B6

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, 5/10, C07K 14/47, 14/705, 16/18, 16/28, A61K 38/17, C12Q 1/68, G01N 33/50

A2

(11) International Publication Number: WO 00/60080

(43) International Publication Date: 12 October 2000 (12.10.00)

(21) International Application Number: PCT/US00/09072

(22) International Filing Date: 4 April 2000 (04.04.00)

(30) Priority Data:

60/127,852	5 April 1999 (05.04.99)	US
60/132,647	5 May 1999 (05.05.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/127,852 (CIP)
Filed on	5 April 1999 (05.04.99)
US	60/132,647 (CIP)
Filed on	5 May 1999 (05.05.99)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MOLECULES OF THE IMMUNE SYSTEM

(57) Abstract

The invention provides human immune system molecules (IMOL) and polynucleotides which identify and encode IMOL. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of IMOL.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

MOLECULES OF THE IMMUNE SYSTEM

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of immune system molecules and to the use of these sequences in the diagnosis, treatment, and prevention of immunological disorders, infections, and cell proliferative disorders including cancer.

BACKGROUND OF THE INVENTION

All vertebrates have developed sophisticated and complex immune systems that provide protection from viral, bacterial, fungal and parasitic infections and cancers. Included in these systems are the processes of humoral immunity, the complement cascade, and the inflammatory response (See Paul, W.E. (1993) Fundamental Immunology, Raven Press, Ltd., New York NY pp.1-20).

The immune system responds to invading microorganisms in two major ways: antibody production and cell mediated responses. The cells that recoginze and destroy pathogens include different types of leukocytes: monocytes, lymphocytes, neutrophils, eosinophils, and basophils. Neutrophils and monocytes attack invading bacteria, viruses, and other pathogens and destroy them by phagocytosis. Monocytes enter tissues and differentiate into macrophages which are extremely phagocytic. Lymphocytes and plasma cells are a part of the immune system which recognizes specific foreign molecules and organisms and inactivates them, as well as signaling other cells to attack the invaders. Leukocytes are formed from two stem cell lineages in bone marrow. The myeloid stem cell line produces granulocytes and monocytes and, the lymphoid stem cell produces lymphocytes. Lymphoid cells travel to the thymus, spleen and lymph nodes, where they mature and differentiate into lymphocytes. Two classes of lymphocytes are T- and B- lymphocytes, also called T cells and B cells. The maturation of lymphocytes is subject to control by a variety of factors including the interleukins.

Cell-mediated immune responses involve T cells that react with foreign antigen on the surface of infected or transformed (cancerous) cells. There are two major types of T cells: cytotoxic T cells destroy antigen-bearing cells, whereas helper T cells activate other white blood cells via chemical signals such as the interleukins. One class of helper cell, T_H1, activates macrophages to destroy ingested microorganisms, while another, T_H2, stimulates the production of antibodies by B cells.

Antibodies are immunoglobulin proteins produced by B-lymphocytes which bind to specific antigens and cause inactivation or promote destruction of the antigen by other cells. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two

identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ϵ, γ, and μ H-chain types. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp.1206-1213 and 1216-1217.)

Both H-chains and L-chains contain repeated immunoglobulin domains. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region. In addition, H chains such as μ have been shown to associate with other polypeptides during differentiation of the B-cell.

Host defense is further augmented by the complement system. The complement system serves as an effector system and is involved in infectious agent recognition. It can function as an independent immune network or in conjunction with other humoral immune responses. The complement system is comprised of numerous plasma and membrane proteins that act in a cascade of reaction sequences whereby one component activates the next to generate a rapid and amplified response to infection through either an inflammatory response or increased phagocytosis. The first element in this cascade, C1, is composed of subunits C1q, C1r, and C1s. C1q-induced cellular responses are thought to be involved in host defense and in protection against autoimmunity since C1q-deficient humans are suceptible to infectious disorders and a very high incidence of autoimmune disease (Nicholson-Weller, A. and Klickstein, L.B. (1999) Curr. Opin. Immunol. 11:42-46).

The classical pathway requires antibody binding to infectious agent antigens. The antibodies serve to define the target and initiate the complement system cascade, culminating in the destruction

of the infectious agent. In this pathway, since the antibody guides initiation of the process, the complement can be seen as an effector arm of the humoral immune system.

The alternative pathway of the complement system does not require the presence of pre-existing antibodies for targeting infectious agent destruction. Rather, this pathway, through low levels of an activated component, remains constantly primed and provides surveillance in the non-immune host to enable targeting and destruction of infectious agents. In this case foreign material triggers the cascade, thereby facilitating phagocytosis or lysis (Paul, W. (1993) Fundamental Immunology Raven Press Ltd., New York NY pp.918-919).

Antigen Recognition

A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include T-cell receptors, major histocompatibility (MHC) proteins, antibodies, and immune cell-specific surface markers such as CD4, CD8, and CD28.

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246.)

Antibodies, or immunoglobulins, bind and neutralize antigens in the circulation and other extracellular fluids. Antibody classes include IgG, IgA, IgM, IgD, and IgE. IgA are found primarily in secretions and play an important role in mucosal immunity. IgA are transcytosed across epithelial cell sheets and secreted along with mucous into the lumenal space. Transcytosis and secretion are

mediated by the polymeric Ig receptor, which binds to and transports IgA. The polymeric Ig receptor is a transmembrane protein with about two to five Ig domains in its extracellular domain. (Reviewed in Alberts, supra, pp. 1210-1211; Kulseth, M. A. et al. (1995) DNA Cell Biol. (1995) 14:251-256.)

T-cell receptors are both structurally and functionally related to antibodies. T-cell receptors are cell surface proteins that bind foreign antigens and mediate diverse aspects of the immune response. A typical T-cell receptor is a heterodimer comprised of two disulfide-linked polypeptide chains called α and β. Each chain is about 280 amino acids in length and contains one variable region and one constant region. Each variable or constant region folds into an Ig domain. The variable regions from the α and β chains come together in the heterodimer to form the antigen recognition site. T-cell receptor diversity is generated by somatic rearrangement of gene segments encoding the α and β chains. T-cell receptors recognize small peptide antigens that are expressed on the surface of antigen-presenting cells and pathogen-infected cells. These peptide antigens are presented on the cell surface in association with major histocompatibility proteins which provide the proper context for antigen recognition. (Reviewed in Alberts, supra, pp. 1228-1229.)

Immune Cell Signaling

Cytokines comprise a family of signaling molecules that modulate the immune system and the inflammatory response. Cytokines are usually secreted by leukocytes, or white blood cells, in response to injury or infection. However, other tissues are capable of secreting cytokines in response to disease or trauma. Cytokines function as growth and differentiation factors that act primarily on cells of the immune system including B- and T-lymphocytes, monocytes, macrophages, granulocytes, and their progenitors, such as the myeloid stem cells and lymphoid stem cells. Like other signaling molecules, cytokines bind to specific plasma membrane receptors and trigger intracellular signal transduction pathways which alter gene expression patterns. There is considerable potential for the use of cytokines in the treatment of inflammation and immune system disorders.

Cytokine structure and function have been extensively characterized in vitro. Most cytokines are small polypeptides of about 30 kilodaltons or less. Over 50 cytokines have been identified from human and rodent sources. Examples of cytokine subfamilies include the interferons (IFN-α, -β, and -γ), the interleukins (IL-1 through IL-18), the tumor necrosis factors (TNF-α and -β), and the chemokines. Many cytokines have been produced using recombinant DNA techniques, and the activities of individual cytokines have been determined in vitro. These activities include regulation of leukocyte proliferation, differentiation, and motility. (Reviewed in Callard, R. E. and Gearing, A. J. H. (1994) The Cytokine Facts Book, Academic Press, San Diego, CA, pp. 2-6, 12-17; see, for example, Fossiez, F. et al. (1998) Int. Rev. Immunol. 16:541-551.)

The activity of an individual cytokine in vitro may not reflect the full scope of that cytokine's

activity in vivo. Cytokines are not expressed individually in vivo but are instead expressed in combination with a multitude of other cytokines when the organism is challenged with a stimulus. Together, these cytokines collectively modulate the immune response in a manner appropriate for that particular stimulus. Therefore, the physiological activity of a cytokine is determined by the stimulus itself and by complex interactive networks among co-expressed cytokines which may demonstrate both synergistic and antagonistic relationships.

Chemokines comprise a cytokine subfamily with over 30 members. (Reviewed in Wells, T. N. C. and Peitsch, M. C. (1997) J. Leukoc. Biol. 61:545-550.) Chemokines were initially identified as chemotactic proteins that recruit monocytes and macrophages to sites of inflammation. Recent evidence indicates that chemokines may also play key roles in hematopoiesis and HIV-1 infection. Chemokines are small proteins which range from about 6-15 kilodaltons in molecular weight. Chemokines are further classified as C, CC, CXC, or CX_3C based on the number and position of critical cysteine residues. The CXC chemokines, for example, each contain a conserved motif consisting of two cysteines separated by a single residue followed by two additional cysteines which occur downstream at about 23- and 12-residue intervals (ExPASy PROSITE database, documents PS00471 and PDOC00434). The presence and spacing of these four cysteine residues are highly conserved, whereas the intervening residues diverge significantly.

Organs of the Immune System

The major organs of the immune system are classified as either primary or secondary lymphoid organs. Primary lymphoid organs include the bone marrow, which produces B-cells, and the thymus, which produces T-cells (thymocytes). Bone marrow also contains blood vessels, nerves, fatty tissue, and stromal cells. Stromal cells produce a supporting meshwork of collagen fibers and other extracellular matrix components which are important for promoting the growth and differentiation of B-cells and other hematopoietic cells. Upon maturation, B- and T-cells travel through the lymphatic system and populate secondary lymphoid organs throughout the body such as the lymph nodes, adenoids, tonsils, spleen, and intestinal Peyer's patches.

Disorders of the Immune System

Disorders of the immune system include various autoimmune and inflammatory diseases caused by failure of the immune system to discriminate self from non-self molecules. Other immune system disorders are caused by uncontrolled cell proliferation, including leukemias such as multiple myeloma and lymphomas such as Hodgkin's disease. Immunodeficiency, brought on by a variety of diseases and agents including HIV, renders afflicted individuals susceptible to severe and sometimes fatal bacterial and viral infections. (See, for example, Golub, E. S. et al. (1987) Immunology: A

Synthesis, Sinauer Associates, Sunderland, MA, pages 481 and 509-530.)

Diseases which cause over- or under-abundance of any one type of leukocyte usually result in the entire immune defense system becoming involved. The most well known autoimmune disease is AIDS (Acquired Immunodeficiency Syndrome). This disease depletes the number of helper T cells and leaves the patient susceptible to infection by microorganisms and parasites. Immunocompromised patients are also at increased risk for cancers.

Leukopenia or agranulocytosis occurs when the bone marrow stops producing white blood cells. This leaves the body unprotected against foreign microorganisms, including those which normally inhabit the skin, mucous membranes, and gastrointestinal tract. Impaired phagocytosis occurs in several diseases, including monocytic leukemia, systemic lupus, and granulomatous disease.

Leukemias are an excess production of white blood cells, to the point where a major portion of the body's metabolic resources are directed solely at proliferation of white blood cells, leaving other tissues to starve. In myelogenous leukemias, cancerous young myelogenous cells spread from the bone marrow to other organs, especially the spleen, liver, lymph nodes and other highly vascularized regions. Usually, the extra leukemic cells released are immature, incapable of function, and undifferentiated. Leukemias may be caused by exposure to environmental factors such as radiation or toxic chemicals or by genetic aberration.

Transplant rejection and allergies are examples of situations in which it may be desirable to curtail the immune response. Interleukin activity may play a pivotal role in specific immune responses. An antagonist to IL-17 has been shown to promote survival in heart tissue grafts (Antonysamy, M.A. et al. (1999) Transplant. Proc. 31:93).

The discovery of new immune system molecules and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of immunological disorders, infections, and cell proliferative disorders including cancer.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, immune system molecules, referred to collectively as "IMOL" and individually as "IMOL-1," "IMOL-2," "IMOL-3," "IMOL-4," "IMOL-5," "IMOL-6" "IMOL-7," "IMOL-8," "IMOL-9," "IMOL-10," "IMOL-11," "IMOL-12," "IMOL-13," "IMOL-14," and "IMOL-15." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an

immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-15.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:16-30.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional IMOL, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) exposing a sample comprising the polypeptide to a

compound. and b) detecting agonist activity in the sample. In one alternative. the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional IMOL. comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional IMOL, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:16-30, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding IMOL.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of IMOL.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones

encoding IMOL were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze IMOL, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"IMOL" refers to the amino acid sequences of substantially purified IMOL obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of IMOL. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of IMOL either by directly interacting with IMOL or by acting on components of the biological pathway in which IMOL participates.

An "allelic variant" is an alternative form of the gene encoding IMOL. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to

allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding IMOL include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as IMOL or a polypeptide with at least one functional characteristic of IMOL. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding IMOL, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding IMOL. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent IMOL. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of IMOL is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of IMOL. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of IMOL either by directly interacting with IMOL or by acting on components of the biological pathway in which IMOL participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant.

Antibodies that bind IMOL polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic IMOL, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid

(PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding IMOL or fragments of IMOL may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide

backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of IMOL or the polynucleotide encoding IMOL which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:16-30 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:16-30, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:16-30 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:16-30 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:16-30 and the region of SEQ ID NO:16-30 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-15 is encoded by a fragment of SEQ ID NO:16-30. A fragment of SEQ ID NO:1-15 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-15. For example, a fragment of SEQ ID NO:1-15 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-15. The precise length of

a fragment of SEQ ID NO:1-15 and the region of SEQ ID NO:1-15 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available

from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific

binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression

of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of IMOL which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of IMOL which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of IMOL. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of IMOL.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding IMOL, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purp se such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding IMOL, or fragments thereof, or IMOL itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,

microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in th art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to

another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human immune system molecules (IMOL), the polynucleotides encoding IMOL, and the use of these compositions for the diagnosis, treatment, or prevention of immunological disorders, infections, and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding IMOL. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each IMOL were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each IMOL and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding IMOL. The first column of Table 3 lists the

nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:16-30 and to distinguish between SEQ ID NO:16-30 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express IMOL as a fraction of total tissues expressing IMOL. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing IMOL as a fraction of total tissues expressing IMOL. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding IMOL were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:26 maps to chromosome 16 within the interval from 19.70 to 33.30 centiMorgans. This interval also contains genes and ESTs associated with B cell maturation and MHC Class II transactivation. SEQ ID NO:29 maps to chromosome 11 within the interval from 104.80 to 123.50 centiMorgans. This interval also contains genes and ESTs associated with human lymphoma.

The invention also encompasses IMOL variants. A preferred IMOL variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the IMOL amino acid sequence, and which contains at least one functional or structural characteristic of IMOL.

The invention also encompasses polynucleotides which encode IMOL. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:16-30, which encodes IMOL. The polynucleotide sequences of SEQ ID NO:16-30, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding IMOL. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding IMOL. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:16-30 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:16-30. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of IMOL.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code. a multitude of polynucleotide sequences encoding IMOL, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring IMOL, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode IMOL and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring IMOL under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding IMOL or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding IMOL and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode IMOL and IMOL derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding IMOL or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:16-30 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is

automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton. Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding IMOL may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze

the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode IMOL may be cloned in recombinant DNA molecules that direct expression of IMOL, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express IMOL.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter IMOL-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of IMOL, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable

manner.

In another embodiment, sequences encoding IMOL may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, IMOL itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of IMOL, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active IMOL, the nucleotide sequences encoding IMOL or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding IMOL. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding IMOL. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding IMOL and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding IMOL and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A

Laboratory Manual, Cold Spring Harbor Press, Plainview NY. ch. 4, 8, and 16-17: Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons. New York NY. ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding IMOL. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding IMOL. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding IMOL can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding IMOL into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of IMOL are needed, e.g. for the production of antibodies, vectors which direct high level expression of IMOL may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of IMOL. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of IMOL. Transcription of sequences encoding IMOL may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.)

These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding IMOL may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses IMOL in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of IMOL in cell lines is preferred. For example, sequences encoding IMOL can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc.

Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding IMOL is inserted within a marker gene sequence, transformed cells containing sequences encoding IMOL can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding IMOL under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding IMOL and that express IMOL may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of IMOL using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on IMOL is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding IMOL include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding IMOL, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety

of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding IMOL may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode IMOL may be designed to contain signal sequences which direct secretion of IMOL through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding IMOL may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric IMOL protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of IMOL activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the IMOL encoding sequence and the heterologous protein sequence, so that IMOL may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10).

A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled IMOL may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of IMOL may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of IMOL may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of IMOL and immune system molecules. In addition, the expression of IMOL is closely associated with immunological disorders, infections, and cell proliferative disorders including cancer. SEQ ID NO:26 maps to a chromosomal interval which also contains genes and ESTs associated with B cell maturation and MHC Class II transactivation. SEQ ID NO:29 maps to a chromosomal interval which also contains genes and ESTs associated with human lymphoma. Therefore, IMOL appears to play a role in immunological disorders, infections, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased IMOL expression or activity, it is desirable to decrease the expression or activity of IMOL. In the treatment of disorders associated with decreased IMOL expression or activity, it is desirable to increase the expression or activity of IMOL.

Therefore, in one embodiment, IMOL or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IMOL. Examples of such disorders include, but are not limited to, an immunological disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel

syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other mycosis-causing fungal agent; and an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematode such as trichinella, intestinal nematode such as ascaris, lymphatic filarial nematode, trematode such as schistosoma, and cestode such as tapeworm; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing IMOL or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IMOL including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified IMOL in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of IMOL including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of IMOL may be administered to a subject to treat or prevent a dis rder associated with decreased expression or activity of IMOL including, but not limited to, those listed above.

In a further embodiment, an antagonist of IMOL may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of IMOL. Examples of such disorders include, but are not limited to, those immunological disorders, infections, and cell proliferative disorders including cancer, listed above. In one aspect, an antibody which specifically binds IMOL may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express IMOL.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding IMOL may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of IMOL including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of IMOL may be produced using methods which are generally known in the art. In particular, purified IMOL may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind IMOL. Antibodies to IMOL may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with IMOL or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to IMOL have an amino acid sequence consisting of at least about 5 amino acids, and generally will

consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of IMOL amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to IMOL may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce IMOL-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for IMOL may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between IMOL and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies

reactive to two non-interfering IMOL epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for IMOL. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of IMOL-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple IMOL epitopes, represents the average affinity, or avidity, of the antibodies for IMOL. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular IMOL epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the IMOL-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of IMOL, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of IMOL-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding IMOL, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding IMOL may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding IMOL. Thus, complementary molecules or fragments may be used to modulate IMOL activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding IMOL.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted

organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding IMOL. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding IMOL can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding IMOL. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding IMOL. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding IMOL.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques

for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding IMOL. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of IMOL, antibodies to IMOL, and mimetics, agonists, antagonists, or inhibitors of IMOL. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil,

or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of IMOL, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example IMOL or fragments thereof, antibodies of IMOL, and agonists, antagonists or inhibitors of IMOL, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such

compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind IMOL may be used for the diagnosis of disorders characterized by expression of IMOL, or in assays to monitor patients being treated with IMOL or agonists, antagonists, or inhibitors of IMOL. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for IMOL include methods which utilize the antibody and a label to detect IMOL in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring IMOL, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of IMOL expression. Normal or standard values for IMOL expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to IMOL under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of IMOL expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding IMOL may be used for

diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of IMOL may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of IMOL, and to monitor regulation of IMOL levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding IMOL or closely related molecules may be used to identify nucleic acid sequences which encode IMOL. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding IMOL, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the IMOL encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:16-30 or from genomic sequences including promoters, enhancers, and introns of the IMOL gene.

Means for producing specific hybridization probes for DNAs encoding IMOL include the cloning of polynucleotide sequences encoding IMOL or IMOL derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes <u>in vitro</u> by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding IMOL may be used for the diagnosis of disorders associated with expression of IMOL. Examples of such disorders include, but are not limited to, an immunological disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation,

myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and an infection caused by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, or togavirus; an infection caused by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, or campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, or mycoplasma; an infection caused by a fungal agent classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, or other mycosis-causing fungal agent; and an infection caused by a parasite classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematode such as trichinella, intestinal nematode such as ascaris, lymphatic filarial nematode, trematode such as schistosoma, and cestrode such as tapeworm; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding IMOL may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered IMOL expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding IMOL may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding IMOL may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a

standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding IMOL in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of IMOL, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding IMOL, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding IMOL may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding IMOL, or a fragment of a polynucleotide complementary to the polynucleotide encoding IMOL, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of IMOL include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C.

et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding IMOL may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding IMOL on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides

valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, IMOL, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between IMOL and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with IMOL, or fragments thereof, and washed. Bound IMOL is then detected by methods well known in the art. Purified IMOL can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding IMOL specifically compete with a test compound for binding IMOL. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with IMOL.

In additional embodiments, the nucleotide sequences which encode IMOL may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/127,852 and U.S. Ser. No. 60/132,647, are hereby expressly incorporated by reference.

EXAMPLES

I. Constructi n of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or

without lyophilization. at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:16-30. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although

lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding IMOL occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of IMOL Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:16-30 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:16-30 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data availabl from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location. The genetic map locations of SEQ ID NO:26 and SEQ ID NO:29 are described in the Invention as a range, or interval, of a particular human chromosome. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention.

VI. Extension of IMOL Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:16-30 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30

nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following

parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:16-30 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:16-30 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, <u>supra</u>.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and

patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

IX. Complementary Polynucleotides

Sequences complementary to the IMOL-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring IMOL. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of IMOL. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the IMOL-encoding transcript.

X. Expression of IMOL

Expression and purification of IMOL is achieved using bacterial or virus-based expression systems. For expression of IMOL in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express IMOL upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of IMOL in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding IMOL by either homologous recombination or bacterial-mediated

transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, IMOL is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from IMOL at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified IMOL obtained by these methods can be used directly in the following activity assay.

XI. Demonstration of IMOL Activity

An assay for IMOL activity measures the proliferation of leukocytes in response to IMOL. In this assay, the amount of tritiated thymidine incorporated into newly synthesized DNA is used to estimate proliferative activity. Varying amounts of IMOL are added to cultured leukocytes, such as granulocytes, monocytes, or lymphocytes, in the presence of [^{3}H]thymidine, a radioactive DNA precursor. IMOL for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold IMOL concentration range is indicative of IMOL activity. One unit of activity per milliliter is conventionally defined as the concentration of IMOL producing a 50% response level, where 100% represents maximal incorporation of [^{3}H]thymidine into acid-precipitable DNA.

Alternatively, an assay for IMOL activity measures the expression of IMOL on the cell surface. cDNA encoding IMOL is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M. A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using IMOL-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled

immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of IMOL expressed on the cell surface.

Alternatively, IMOL activity is exemplified by that of immunoglobulins, which recognize and precipitate antigens from serum. The quantitative precipitin reaction measures this activity. (Golub, E. S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.) IMOL is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled IMOL. IMOL-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable IMOL-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable IMOL-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable IMOL-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable IMOL-antigen complex is a measure of IMOL activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

XII. Functional Assays

IMOL function is assessed by expressing the sequences encoding IMOL at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of

fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of IMOL on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding IMOL and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding IMOL and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of IMOL Specific Antibodies

IMOL substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the IMOL amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-IMOL activity by, for example, binding the peptide or IMOL to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring IMOL Using Specific Antibodies

Naturally occurring or recombinant IMOL is substantially purified by immunoaffinity chromatography using antibodies specific for IMOL. An immunoaffinity column is constructed by covalently coupling anti-IMOL antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing IMOL are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of IMOL (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt

antibody/IMOL binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and IMOL is collected.

XV. Identification of Molecules Which Interact with IMOL

IMOL, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled IMOL, washed, and any wells with labeled IMOL complex are assayed. Data obtained using different concentrations of IMOL are used to calculate values for the number, affinity, and association of IMOL with the candidate molecules.

Alternatively, molecules interacting with IMOL are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	16	2705028	PONSAZT01	659491H1 (BRAINOT03), 749439R1 (BRAITUT01), 786579R6 (PROSNOT05), 913159R6 (STOMNOT02), 1440816F6 (THYRNOT03), 2209623H1 (SINTFET03), 2603630F6 (LUNGTUT07), 2705028H1 (PONSAZT01), 4934049F6 (BRSTTUT20), 4934049T6 (BRSTTUT20)
2	17	2751129	THP1AZS08	1441709F1 (THYRNOT03), 2595386F6 (OVARTUT02), 2751129H1 (THP1AZS08), 2751129R6 (THP1AZS08)
3	18	2818076	BRSTNOT14	1269389T1 (BRAINOT09), 2818076F6 (BRSTNOT14), 2818076H1 (BRSTNOT14), 3224685R6 (UTRSNON03), 3224685T6 (UTRSNON03)
4	19	2907049	THYMNOT05	2907049H1 (THYMNOT05), SBZA00343V1, SCCA03402V1, SBZA05102V1
5	20	3402252	ESOGNOT03	2960277T6 (ADRENOT09), 3402252H1 (ESOGNOT03), SBZA00824V1, SBZA01322V1
6	21	3577142	BRONNOT01	2967181T6 (SCORNOT04), 3577142H1 (BRONNOT01)
7	22	3725986	BRSTNOT23	2296581R6 (BRSTNOT05), 3725986H1 (BRSTNOT23), 3957555H2 (HEARFET02), 4325138H1 (TLYMUNT01)
8	23	3799011	SPLNNOT12	1674311F6 (BLADNOT05), 1994530H1 (BRSTTUT03), 3054149H1 (LNODNOT08), 3339240H1 (SPLNNOT10)
9	24	3887384	UTRSNOT05	078084F1 (SYNORAB01), 3887384H1 (UTRSNOT05), 5422664H1 (PROSTMT07)
10	25	1352789	LATRTUT02	530695T6 (BRAINOT03), 1352789H1 and 1352789T6 (LATRTUT02), 1512934F6 and 1512934T6 (PANCTUT01)
11	26	1666486	BMARNOT03	689749R6 (LUNGTUT02), 1666486H1 (BMARNOT03), 2771470F6.comp (COLANOT02)
12	27	1706596	DUODNOT02	077057R1 (SYNORAB01), 1241315R6 (LUNGNOT03), 1370244T6.comp (BSTMNON02), 1706596H1 (DUODNOT02)
13	28	1890540	BLADTUT07	137725F1 and 137725R1 (SYNORAB01), 1890540H1 (BLADTUT07), 2861361F6 (SININOT03), 5370394H1 (BRAINOT22)
14	29	2774913	PANCNOT15	961128R2 (BRSTTUT03), 1511626F1 (LUNGNOT14), 2774913H1 (PANCNOT15), 3295703H1 (TLYJINT01), SAXB00207F1, SAXB00564F1, SASA00067F1
15	30	5571291	TLYMNOT08	534078T6 (BRAINOT03), 1655166F6 (PROSTUT08), 2073461H1 (ISLTNOT01), 2545752F6 (UTRSNOT11), 3171142T6 (BRSTNOT18), 3767674T6 (BRSTNOT24), 3842267H1 and 3845101H1 (DENDNOT01), 5571291H1 (TLYMNOT08)

Table 2

Polyp ptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
1	613	T52 S80 T175 S205 T215 T271 S141 T192 T250 T388 T403 S424 S465 S526 S562 Y404	N50 N327 N463	Polymeric Ig receptor signatures: T250-G281, Y374-E421 Ig domains: G42-T129, G179-A272 G319-A408, G455-P546 Signal peptide: M1-A27 Transmembrane domain: P583-M607	Ig-like membrane protein (g3676136)	BLOCKS PFAM HMM MOTIFS BLAST SPSCAN
2	271	T5 S19 S21 T89 T208 S11 T63 T68 S82 S180 S187 S192 S193 S203 T225 T251			NY-REN-45 antigen Associated with renal cell carcinoma (g5360115)	MOTIFS BLAST
3	235	T43 T51 S93 S105 T230 S79 S113			IL-6-induced myeloid differentiation (MyD88) protein (g53294)	MOTIFS BLAST
4	310	S308 T129 S148 S235 S85 T112 T155 S208	N89 N201	Ig domain: G35-T112 Signal peptide: M1-A21 Transmembrane domain: Y280-M300	T-cell receptor beta (g1100182)	PFAM HMM MOTIFS BLAST SPSCAN
5	246	T114 S186 S203 S153	N68 N82	MHC I domains: D29-D65, T120-L205 Signal peptide: M1-R27	MHC class I antigen (g2789617)	PFAM HMM MOTIFS BLAST SPSCAN
6	180	T131 S24 S27 S93	N75	Small CXC cytokine signature: R120-N128 Signal peptide: M1-P22	cytokine homolog CYT07 (g6013321)	HMM PRINTS MOTIFS BLAST SPSCAN
7	200	T21 S86 S96 S100 T189 S11 S12 S34			B-G (MHC) antigen (g211254)	MOTIFS BLAST

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
8	211	S76 T94 S166 S200 S205	N83 N92	Ig domain: G29-V105 Signal peptide: M1-T23	CMRF-35 antigen (homologous to the poly-Ig receptor) (g396170)	PFAM HMM MOTIFS BLAST SPSCAN
9	225	S12 T23 S61 S65 S76 T88 T138 S150 S151 S159 T209 S218 S14 S19 Y114	N52		Thymocyte protein cThy28kD (g995778)	MOTIFS BLAST
10	329	S74 S75 T99 T127 T154 T194 T299 T66 S75 T134 T222 Y221		ATP/GTP binding site: G37-S44	AIG1 Antivirulence-induced gene protein g1127804	Motifs BLAST
11	237	S36 S74 S89 T125 S185 T187 S205 T44		Ig/major histocompatability complex: D193-E236 Immunoglobulin domains: G38-Q112 S150-V219 Signal peptide: M1-C22	g2765423 Immunoglobulin kappa light chain	Motifs ProfileScan Pfam SPScan HMM BLOCKS BLOCKS-DOMO BLAST
12	235	S109 S144 S223 S68 S79 S210		Ig/major histocompatability complex: N193-S235 Immunoglobulin: R34-V110, A150-V218 Signal Peptide: M1-A21	Immunoglobulin light chain lambda precursor (AA -20 to 215) g33395	Motifs ProfileScan Pfam SPScan HMM BLOCKS BLOCKS-DOMO BLAST
13	246	T77 S138 S48 T152		C1q: A119-L242 C1q domain proteins: G54-E80 G134-M169 D202-R221 S235-E244 Signal Peptide: M1-C22	C1q-related factor g3747097	Motifs Pfam SPScan HMM BLOCKS PRINTS BLAST

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
14	322	S28 S97 Y168	N66	Transmembrane domains: W205-L223, L240-W258 W67-V85 Signal peptide: M1-A50	myeloid upregulated protein g2463265	Motifs HMM SPScan BLAST
15	191	S11 S91 S103 T142 T73			IL-17 receptor g2826476	Motifs BLAST

Table 3

Nucleotide SEQ ID NO:	Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
16	217-261	Nervous (0.377) Reproductive (0.203) Cardiovascular (0.145)	Cancer (0.522) Inflammation (0.246) Trauma (0.101)	pINCY
17	109-153 757-801	Reproductive (0.400) Developmental (0.200) Endocrine (0.200)	Cancer (0.600) Cell Proliferation (0.400)	PSPORT
18	812-856	Reproductive (0.375) Cardiovascular (0.125) Developmental (0.125)	Cancer (0.500) Cell Proliferation (0.250) Trauma (0.250)	pINCY
19	165-209	Hematopoietic/Immune (0.294) Gastrointestinal (0.235) Reproductive (0.132)	Inflammation (0.412) Cancer (0.353) Cell Proliferation (0.132)	pINCY
20	650-694	Cardiovascular (0.250) Gastrointestinal (0.250) Musculoskeletal (0.250)	Cancer (0.250) Cell Proliferation (0.250) Inflammation (0.250)	pINCY
21	541-585	Cardiovascular (0.333) Nervous (0.333) Reproductive (0.333)	Cancer (0.667) Inflammation (0.333)	pINCY
22	165-209 597-641	Nervous (0.400) Reproductive (0.400) Developmental (0.200)	Cancer (0.200) Cell Proliferation (0.200) Inflammation (0.200)	pINCY
23	595-639	Hematopoietic/Immune (0.381) Reproductive (0.190) Cardiovascular (0.143)	Inflammation (0.524) Cancer (0.333) Cell Proliferation (0.143)	pINCY
24	704-748	Reproductive (0.233) Gastrointestinal (0.183) Nervous (0.167)	Cancer (0.483) Inflammation (0.233) Cell Proliferation (0.133)	pINCY
25	84-128	Reproductive (0.194) Cardiovascular (0.164) Hematopoietic/Immune (0.149) Nervous (0.149)	Cancer (0.403) Inflammation (0.283) Fetal/Proliferative/Cell Line (0.075) Trauma (0.104)	pINCY
26	348-392	Reproductive (0.268) Gastrointestinal (0.254) Cardiovascular (0.136)	Cancer (0.573) Inflammation (0.251) Trauma (0.071)	pINCY
27	111-155	Gastrointestinal (0.296) Reproductive (0.251) Cardiovascular (0.133)	Cancer (0.557) Inflammation (.300) Trauma (0.089)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
28	217-261	Developmental (0.162) Musculoskeletal (0.162) Nervous (0.162)	Cancer (0.486) Fetal/Proliferative/Cell Line (0.216) Inflammation (0.162)	pINCY
29	390-434	Reproductive (0.311) Gastrointestinal (0.174) Hematopoietic/Immune (0.121) Cardiovascular (0.121)	Cancer (0.530) Inflammation (0.174) Fetal/Proliferative/Cell Line (0.106)	pINCY
30	226-270	Gastrointestinal (0.214) Hematopoietic/Immune (0.178) Nervous (0.178) Reproductive (0.178)	Cancer (0.571) Inflammation (0.143) Neurological (0.107)	pINCY

Table 4

Polynucleotide SEQ ID NO:	Library	Library Comment
16	PONSAZT01	This library was constructed using RNA isolated from diseased pons tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
17	THP1AZS08	This subtracted library was constructed using 5.76 million clones from a 5-aza-2'-deoxycytidine (AZ)-treated THP-1 (promonocyte) cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The hybridization probes for subtraction were derived from RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., Nuc. Acids Res. (1991) 19:1954, and Bonaldo et al., Genome Research (1996) 6:791. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer 26 (1980):171).
18	BRSTNOT14	This library was constructed using RNA isolated from breast tissue removed from a 62-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Ductal carcinoma in situ, comedo type, comprised 60% of the tumor mass. Metastatic adenocarcinoma was identified in one (of 14) axillary lymph nodes with no perinodal extension. The tumor cells were strongly positive for estrogen receptors and weakly positive for progesterone receptors. Patient history included a benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.
19	THYMNOT05	This library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during a thymectomy and closure of a patent ductus arteriosus. Patient history included cardiac catheterization, echocardiogram, Blalock-Taussig shunt, and pulmonary valvotomy. Family history included benign hypertension, osteoarthritis, depressive disorder, and extrinsic asthma.
20	ESOGNOT03	This library was constructed using RNA isolated from esophageal tissue obtained from a 53-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, and regional lymph node biopsy. Patient history included membranous nephritis, hyperlipidemia, benign hypertension, anxiety state, and adenotonsillectomy. Family history included atherosclerotic coronary artery disease, abdominal aortic aneurysm rupture, and breast cancer.
21	BRONNOT01	This library was constructed using RNA isolated from bronchial tissue removed from a 15-year-old Caucasian male.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
22	BRSTNOT23	This library was constructed using RNA isolated from diseased breast tissue removed from a 35-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included type II diabetes, atherosclerotic coronary artery disease, acute myocardial infarction, hyperlipidemia, and coronary artery bypass.
23	SPLNNOT12	This library was constructed using RNA isolated from spleen tissue removed from a 65-year-old female. Pathology for the associated tumor tissue indicated well-differentiated neuroendocrine carcinoma (islet cell tumor), nuclear grade 1, forming a dominant mass in the distal pancreas. Multiple smaller tumor nodules were immediately adjacent to the main mass. The liver showed metastatic grade 1 islet cell tumor, forming multiple nodules. Multiple (4) pericholedochal lymph nodes contained metastatic grade 1 islet cell tumor.
24	UTRSNOT05	This library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
25	LATRTUT02	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, and hyperlipidemia. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
26	BMARNOT03	Library was constructed using RNA isolated from the left tibial bone marrow tissue of a 16-year-old Caucasian male during a partial left tibial ostectomy with free skin graft. Patient history included an abnormality of the red blood cells. Family history included osteoarthritis.
27	DUODNOT02	Library was constructed using RNA isolated from duodenal tissue of a 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
28	BLADTUT07	Library was constructed using RNA isolated from bladder tumor tissue removed from the anterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated a grade 3 transitional cell carcinoma in the left lateral bladder. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
29	PANCNOT15	Library was constructed using RNA isolated from diseased pancreatic tissue removed from a 15-year-old Caucasian male during a exploratory laparotomy with distal pancreatectomy and total splenectomy. Pathology indicated islet cell hyperplasia. Family history included prostate cancer and cardiovacular disease.
30	TLYMNOT08	Library was constructed using RNA isolated from anergic allogenic T-lymphocyte tissue removed from an adult (40-50-year-old) Caucasian male. The cells were incubated for 3 days in the presence of OKT3 mAb (1 microgram/ml OKT3) and 5% human serum.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-15,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-15.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:16-30.

5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

6. A cell transformed with a recombinant polynucleotide of claim 5.

7. A transgenic organism comprising a recombinant polynucleotide of claim 5.

8. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

9. An isolated antibody which specifically binds to a polypeptide of claim 1.

10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30,

b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.

12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:

a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.

14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.

15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

16. A method for treating a disease or condition associated with decreased expression of functional IMOL, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.

17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

19. A method for treating a disease or condition associated with decreased expression of functional IMOL, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with overexpression of functional IMOL, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
YUE, Henry
LAL, Preeti
TANG, Y. Tom
BAUGHN, Mariah R.
AZIMZAI, Yalda
LU, Dyung Aina M.

<120> MOLECULES OF THE IMMUNE SYSTEM

<130> PF-0680 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/127,852; 60/132,647
<151> 1999-04-05; 1999-05-05

<160> 30

<170> PERL Program

<210> 1
<211> 613
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte 2705028CD1

<400> 1
Met Gly Ala Leu Arg Pro Thr Leu Leu Pro Pro Ser Leu Pro Leu
1 5 10 15
Leu Leu Leu Leu Met Leu Gly Met Gly Cys Trp Ala Arg Glu Val
20 25 30
Leu Val Pro Glu Gly Pro Leu Tyr Arg Val Ala Gly Thr Ala Val
35 40 45
Ser Ile Ser Cys Asn Val Thr Gly Tyr Glu Gly Pro Ala Gln Gln
50 55 60
Asn Phe Glu Trp Phe Leu Tyr Arg Pro Glu Ala Pro Asp Thr Ala
65 70 75
Leu Gly Ile Val Ser Thr Lys Asp Thr Gln Phe Ser Tyr Ala Val
80 85 90
Phe Lys Ser Arg Val Val Ala Gly Glu Val Gln Val Gln Arg Leu
95 100 105
Gln Gly Asp Ala Val Val Leu Lys Ile Ala Arg Leu Gln Ala Gln
110 115 120
Asp Ala Gly Ile Tyr Glu Cys His Thr Pro Ser Thr Asp Thr Arg
125 130 135
Tyr Leu Gly Ser Tyr Ser Gly Lys Val Glu Leu Arg Val Leu Pro
140 145 150
Asp Val Leu Gln Val Ser Ala Ala Pro Pro Gly Pro Arg Gly Arg
155 160 165
Gln Ala Pro Thr Ser Pro Pro Arg Met Thr Val His Glu Gly Gln
170 175 180
Glu Leu Ala Leu Gly Cys Leu Ala Arg Thr Ser Thr Gln Lys His
185 190 195
Thr His Leu Ala Val Ser Phe Gly Arg Ser Val Pro Glu Ala Pro
200 205 210
Val Gly Arg Ser Thr Leu Gln Glu Val Val Gly Ile Arg Ser Asp

```
                215                     220                     225
Leu Ala Val Glu Ala Gly Ala Pro Tyr Ala Glu Arg Leu Ala Ala
                230                     235                     240
Gly Glu Leu Arg Leu Gly Lys Glu Gly Thr Asp Arg Tyr Arg Met
                245                     250                     255
Val Val Gly Gly Ala Gln Ala Gly Asp Ala Gly Thr Tyr His Cys
                260                     265                     270
Thr Ala Ala Glu Trp Ile Gln Asp Pro Asp Gly Ser Trp Ala Gln
                275                     280                     285
Ile Ala Glu Lys Arg Ala Val Leu Ala His Val Asp Val Gln Thr
                290                     295                     300
Leu Ser Ser Gln Leu Ala Val Thr Val Gly Pro Gly Glu Arg Arg
                305                     310                     315
Ile Gly Pro Gly Glu Pro Leu Glu Leu Leu Cys Asn Val Ser Gly
                320                     325                     330
Ala Leu Pro Pro Ala Gly Arg His Ala Ala Tyr Ser Val Gly Trp
                335                     340                     345
Glu Met Ala Pro Ala Gly Ala Pro Gly Pro Gly Arg Leu Val Ala
                350                     355                     360
Gln Leu Asp Thr Glu Gly Val Gly Ser Leu Gly Pro Gly Tyr Glu
                365                     370                     375
Gly Arg His Ile Ala Met Glu Lys Val Ala Ser Arg Thr Tyr Arg
                380                     385                     390
Leu Arg Leu Glu Ala Ala Arg Pro Gly Asp Ala Gly Thr Tyr Arg
                395                     400                     405
Cys Leu Ala Lys Ala Tyr Val Arg Gly Ser Gly Thr Arg Leu Arg
                410                     415                     420
Glu Ala Ala Ser Ala Arg Ser Arg Pro Leu Pro Val His Val Arg
                425                     430                     435
Glu Glu Gly Val Val Leu Glu Ala Val Ala Trp Leu Ala Gly Gly
                440                     445                     450
Thr Val Tyr Arg Gly Glu Thr Ala Ser Leu Leu Cys Asn Ile Ser
                455                     460                     465
Val Arg Gly Gly Pro Pro Gly Leu Arg Leu Ala Ala Ser Trp Trp
                470                     475                     480
Val Glu Arg Pro Glu Asp Gly Glu Leu Ser Ser Val Pro Ala Gln
                485                     490                     495
Leu Val Gly Gly Val Gly Gln Asp Gly Val Ala Glu Leu Gly Val
                500                     505                     510
Arg Pro Gly Gly Gly Pro Val Ser Val Glu Leu Val Gly Pro Arg
                515                     520                     525
Ser His Arg Leu Arg Leu His Ser Leu Gly Pro Glu Asp Glu Gly
                530                     535                     540
Val Tyr His Cys Ala Pro Ser Ala Trp Val Gln His Ala Asp Tyr
                545                     550                     555
Ser Trp Tyr Gln Ala Gly Ser Ala Arg Ser Gly Pro Val Thr Val
                560                     565                     570
Tyr Pro Tyr Met His Ala Leu Asp Thr Leu Phe Val Pro Leu Leu
                575                     580                     585
Val Gly Thr Gly Val Ala Leu Val Thr Gly Ala Thr Val Leu Gly
                590                     595                     600
Thr Ile Thr Cys Cys Phe Met Lys Arg Leu Arg Lys Arg
                605                     610
```

<210> 2
<211> 271
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No.: 2751129CD1

<400> 2
```
Met Val Val Val Thr Gly Arg Glu Pro Asp Ser Arg Arg Gln Asp
 1               5                  10                  15
Gly Ala Met Ser Ser Ser Asp Ala Glu Asp Asp Phe Leu Glu Pro
                20                  25                  30
Ala Thr Pro Thr Ala Thr Gln Ala Gly His Ala Leu Pro Leu Leu
                35                  40                  45
Pro Gln Glu Phe Pro Glu Val Val Pro Leu Asn Ile Gly Gly Ala
                50                  55                  60
His Phe Thr Thr Arg Leu Ser Thr Leu Arg Cys Tyr Glu Asp Thr
                65                  70                  75
Met Leu Ala Ala Met Phe Ser Gly Arg His Tyr Ile Pro Thr Asp
                80                  85                  90
Ser Glu Gly Arg Tyr Phe Ile Asp Arg Asp Gly Thr His Phe Gly
                95                 100                 105
Asp Val Leu Asn Phe Leu Arg Ser Gly Asp Leu Pro Pro Arg Glu
               110                 115                 120
Arg Val Arg Ala Val Tyr Lys Glu Ala Gln Tyr Tyr Ala Ile Gly
               125                 130                 135
Pro Leu Leu Glu Gln Leu Glu Asn Met Gln Pro Leu Lys Gly Glu
               140                 145                 150
Lys Val Arg Gln Ala Phe Leu Gly Leu Met Pro Tyr Tyr Lys Asp
               155                 160                 165
His Leu Glu Arg Ile Val Glu Ile Ala Gly Cys Val Arg Ser Ser
               170                 175                 180
Gly Arg Pro Ala Leu Pro Ser Ser Arg Ser Val Ser Ser Arg Arg
               185                 190                 195
Arg Cys Pro Ser Pro Pro Met Ser Val Arg Ser Ser Thr Pro Cys
               200                 205                 210
Asp Leu Ser Gly Val Arg Val Thr Gly Ser Phe Leu Ser Thr Thr
               215                 220                 225
Val Lys Trp Met Cys Leu Leu Gly Pro Gly Arg Leu Trp Leu Met
               230                 235                 240
Phe Met Thr Cys Cys Thr Ala Trp Ser Arg Thr Ser Arg Pro Arg
               245                 250                 255
Val Ser Pro Trp Thr Thr Ser Ala Ser Gly Cys Val Thr Ser Thr
               260                 265                 270
Ser
```

<210> 3
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2818076CD1

<400> 3
```
Met Ala Ser Ser Thr Ser Leu Pro Ala Pro Gly Ser Arg Pro Lys
 1               5                  10                  15
Lys Pro Leu Gly Lys Met Ala Asp Trp Phe Arg Gln Thr Leu Leu
                20                  25                  30
Lys Lys Pro Lys Lys Arg Pro Asn Ser Pro Glu Ser Thr Ser Ser
                35                  40                  45
Asp Ala Ser Gln Pro Thr Ser Gln Asp Ser Pro Leu Pro Pro Ser
                50                  55                  60
Leu Ser Ser Val Thr Ser Pro Ser Leu Pro Pro Thr His Ala Ser
                65                  70                  75
```

Asp Ser Gly Ser Ser Arg Trp Ser Lys Asp Tyr Asp Val Cys Val
80 85 90
Cys His Ser Glu Glu Asp Leu Val Ala Ala Gln Asp Leu Val Ser
95 100 105
Tyr Leu Glu Gly Ser Thr Ala Ser Leu Arg Cys Phe Leu Gln Leu
110 115 120
Arg Asp Ala Thr Pro Gly Gly Ala Ile Val Ser Glu Leu Cys Gln
125 130 135
Ala Leu Ser Ser Ser His Cys Arg Val Leu Leu Ile Thr Pro Gly
140 145 150
Phe Leu Gln Asp Pro Trp Cys Lys Tyr Gln Met Leu Gln Ala Leu
155 160 165
Thr Glu Ala Pro Gly Ala Glu Gly Cys Thr Ile Pro Leu Leu Leu
170 175 180
Gly Leu Ser Arg Ala Ala Tyr Pro Pro Glu Leu Arg Phe Met Tyr
185 190 195
Tyr Val Asp Gly Arg Gly Pro Asp Gly Gly Phe Arg Gln Val Lys
200 205 210
Glu Ala Val Met Arg Cys Lys Leu Leu Gln Glu Gly Glu Gly Glu
215 220 225
Arg Asp Ser Ala Thr Val Ser Asp Leu Leu
230 235

<210> 4
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2907049CD1

<400> 4
Met Gly Pro Gly Leu Leu His Trp Met Ala Leu Cys Leu Leu Gly
1 5 10 15
Thr Gly His Gly Asp Ala Met Val Ile Gln Asn Pro Arg Tyr Gln
20 25 30
Val Thr Gln Phe Gly Lys Pro Val Thr Leu Ser Cys Ser Gln Thr
35 40 45
Leu Asn His Asn Val Met Tyr Trp Tyr Gln Gln Lys Ser Ser Gln
50 55 60
Ala Pro Lys Leu Leu Phe His Tyr Tyr Asp Lys Asp Phe Asn Asn
65 70 75
Glu Ala Asp Thr Pro Asp Asn Phe Gln Ser Arg Arg Pro Asn Thr
80 85 90
Ser Phe Cys Phe Leu Asp Ile Arg Ser Pro Gly Leu Gly Asp Ala
95 100 105
Ala Met Tyr Leu Cys Ala Thr Ser Lys Tyr Arg Asp Gly Glu Leu
110 115 120
Phe Phe Gly Glu Gly Ser Arg Leu Thr Val Leu Glu Asp Leu Lys
125 130 135
Asn Val Phe Pro Pro Glu Val Ala Val Phe Glu Pro Ser Glu Ala
140 145 150
Glu Ile Ser His Thr Gln Lys Ala Thr Leu Val Cys Leu Ala Thr
155 160 165
Gly Phe Tyr Pro Asp His Val Glu Leu Ser Trp Trp Val Asn Gly
170 175 180
Lys Glu Val His Ser Gly Val Ser Thr Asp Pro Gln Pro Leu Lys
185 190 195
Glu Gln Pro Ala Leu Asn Asp Ser Arg Tyr Cys Leu Ser Ser Arg
200 205 210

```
Leu Arg Val Ser Ala Thr Phe Trp Gln Asn Pro Arg Asn His Phe
                215                 220                 225
Arg Cys Gln Val Gln Phe Tyr Gly Leu Ser Glu Asn Asp Glu Trp
                230                 235                 240
Thr Gln Asp Arg Ala Lys Pro Val Thr Gln Ile Val Ser Ala Glu
                245                 250                 255
Ala Trp Gly Arg Ala Asp Cys Gly Phe Thr Ser Glu Ser Tyr Gln
                260                 265                 270
Gln Gly Val Leu Ser Ala Thr Ile Leu Tyr Glu Ile Leu Leu Gly
                275                 280                 285
Lys Ala Thr Leu Tyr Ala Val Leu Val Ser Ala Leu Val Leu Met
                290                 295                 300
Ala Met Val Lys Arg Lys Asp Ser Arg Gly
                305                 310


<210> 5
<211> 246
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3402252CD1

<400> 5
Met Ala Ala Ala Ala Ile Pro Ala Leu Leu Leu Cys Leu Pro Leu
 1              5                   10                  15
Leu Phe Leu Leu Phe Gly Trp Ser Arg Ala Arg Arg Asp Asp Pro
                20                  25                  30
His Ser Leu Cys Tyr Asp Ile Thr Val Ile Pro Lys Phe Arg Pro
                35                  40                  45
Gly Pro Arg Trp Cys Ala Val Gln Gly Gln Val Asp Glu Lys Thr
                50                  55                  60
Phe Leu His Tyr Asp Cys Gly Asn Lys Thr Val Thr Pro Val Ser
                65                  70                  75
Pro Leu Gly Lys Lys Leu Asn Val Thr Thr Ala Trp Lys Ala Gln
                80                  85                  90
Asn Pro Val Leu Arg Glu Val Val Asp Ile Leu Thr Glu Gln Leu
                95                  100                 105
Arg Asp Ile Gln Leu Glu Asn Tyr Thr Pro Lys Glu Pro Leu Thr
                110                 115                 120
Leu Gln Ala Arg Met Ser Cys Glu Gln Lys Ala Glu Gly His Ser
                125                 130                 135
Ser Gly Ser Trp Gln Phe Ser Phe Asp Gly Gln Ile Phe Leu Leu
                140                 145                 150
Phe Asp Ser Glu Lys Arg Met Trp Thr Thr Val His Pro Gly Ala
                155                 160                 165
Arg Lys Met Lys Glu Lys Trp Glu Asn Asp Lys Val Val Ala Met
                170                 175                 180
Ser Phe His Tyr Phe Ser Met Gly Asp Cys Ile Gly Trp Leu Glu
                185                 190                 195
Asp Phe Leu Met Gly Met Asp Ser Thr Leu Glu Pro Ser Ala Gly
                200                 205                 210
Ala Pro Leu Ala Met Ser Ser Gly Thr Thr Gln Leu Arg Ala Thr
                215                 220                 225
Ala Thr Thr Leu Ile Leu Cys Cys Leu Leu Ile Ile Leu Pro Cys
                230                 235                 240
Phe Ile Leu Pro Gly Ile
                245
```

<210> 6
<211> 180
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3577142CD1

<400> 6
```
Met Asp Trp Pro His Asn Leu Leu Phe Leu Leu Thr Ile Ser Ile
 1               5                  10                  15
Phe Leu Gly Leu Gly Gln Pro Arg Ser Pro Lys Ser Lys Arg Lys
                20                  25                  30
Gly Gln Gly Arg Pro Gly Pro Leu Val Pro Gly Pro His Gln Val
                35                  40                  45
Pro Leu Asp Leu Val Ser Arg Met Lys Pro Tyr Ala Arg Met Glu
                50                  55                  60
Glu Tyr Glu Arg Asn Ile Glu Glu Met Val Ala Gln Leu Arg Asn
                65                  70                  75
Ser Ser Glu Leu Ala Gln Arg Lys Cys Glu Val Asn Leu Gln Leu
                80                  85                  90
Trp Met Ser Asn Lys Arg Ser Leu Ser Pro Trp Gly Tyr Ser Ile
                95                 100                 105
Asn His Asp Pro Ser Arg Ile Pro Val Asp Leu Pro Glu Ala Arg
               110                 115                 120
Cys Leu Cys Leu Gly Cys Val Asn Pro Phe Thr Met Gln Glu Asp
               125                 130                 135
Arg Ser Met Val Ser Val Pro Val Phe Ser Gln Val Pro Val Arg
               140                 145                 150
Arg Arg Leu Cys Pro Pro Pro Pro Arg Thr Gly Pro Cys Arg Gln
               155                 160                 165
Arg Ala Val Met Glu Thr Ile Ala Val Gly Cys Thr Cys Ile Phe
               170                 175                 180
```

<210> 7
<211> 200
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3725986CD1

<400> 7
```
Met Glu Ser Gln Leu Gln Glu Arg Val Glu Ser Ser Arg Arg Ala
 1               5                  10                  15
Val Ser Gln Ile Val Thr Val Tyr Asp Lys Leu Gln Glu Lys Val
                20                  25                  30
Glu Leu Leu Ser Arg Lys Leu Asn Ser Gly Asp Asn Leu Ile Val
                35                  40                  45
Glu Glu Ala Val Gln Glu Leu Asn Ser Phe Leu Ala Gln Glu Asn
                50                  55                  60
Met Arg Leu Gln Glu Leu Thr Asp Leu Leu Gln Glu Lys His Arg
                65                  70                  75
Thr Met Ser Gln Glu Phe Ser Lys Leu Gln Ser Lys Val Glu Thr
                80                  85                  90
Ala Glu Ser Arg Val Ser Val Leu Glu Ser Met Ile Asp Asp Leu
                95                 100                 105
Gln Trp Asp Ile Asp Lys Ile Arg Lys Arg Glu Gln Arg Leu Asn
```

```
                    110                     115                     120
Arg His Leu Ala Glu Val Leu Glu Arg Val Asn Ser Lys Gly Tyr
                    125                     130                     135
Lys Val Tyr Gly Ala Gly Ser Ser Leu Tyr Gly Gly Thr Ile Thr
                    140                     145                     150
Ile Asn Ala Arg Lys Phe Glu Glu Met Asn Ala Glu Leu Glu Glu
                    155                     160                     165
Asn Lys Glu Leu Ala Gln Asn Arg Leu Cys Glu Leu Glu Lys Leu
                    170                     175                     180
Arg Gln Asp Phe Glu Glu Val Thr Thr Gln Asn Glu Lys Leu Lys
                    185                     190                     195
Val Arg Thr His Pro
                    200


<210> 8
<211> 211
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3799011CD1

<400> 8
Met Trp Leu Pro Trp Ala Leu Leu Leu Leu Trp Val Pro Gly Cys
  1               5                  10                  15
Phe Ala Leu Ser Lys Cys Arg Thr Val Ala Gly Pro Val Gly Gly
                 20                  25                  30
Ser Leu Ser Val Gln Cys Pro Tyr Glu Lys Glu His Arg Thr Leu
                 35                  40                  45
Asn Lys Tyr Trp Cys Arg Pro Pro Gln Ile Phe Leu Cys Asp Lys
                 50                  55                  60
Ile Val Glu Thr Lys Gly Ser Ala Gly Lys Arg Asn Gly Arg Val
                 65                  70                  75
Ser Ile Arg Asp Ser Pro Ala Asn Leu Ser Phe Thr Val Thr Leu
                 80                  85                  90
Glu Asn Leu Thr Glu Glu Asp Ala Gly Thr Tyr Trp Cys Gly Val
                 95                 100                 105
Asp Thr Pro Trp Leu Arg Asp Phe His Asp Pro Val Val Glu Val
                110                 115                 120
Glu Val Ser Val Phe Pro Ala Ser Thr Ser Met Thr Pro Ala Ser
                125                 130                 135
Ile Thr Ala Ala Lys Thr Ser Thr Ile Thr Thr Ala Phe Pro Pro
                140                 145                 150
Val Ser Ser Thr Thr Leu Phe Ala Val Gly Ala Thr His Ser Ala
                155                 160                 165
Ser Ile Gln Glu Glu Thr Glu Glu Val Val Asn Ser Gln Leu Pro
                170                 175                 180
Leu Thr Pro Leu Pro Ala Gly Ile Val Ala Ala Ser Val Gly Gly
                185                 190                 195
Gly Leu Pro Ala Ser Leu Glu Asp Val Ser Glu Met Asp Gln Ser
                200                 205                 210
Trp


<210> 9
<211> 225
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No.: 3887384CD1

<400> 9
Met Ser Arg Pro Arg Lys Arg Leu Ala Gly Thr Ser Gly Ser Asp
1 5 10 15
Lys Gly Leu Ser Gly Lys Arg Thr Lys Thr Glu Asn Ser Gly Glu
20 25 30
Ala Leu Ala Lys Val Glu Asp Ser Asn Pro Gln Lys Thr Ser Ala
35 40 45
Thr Lys Asn Cys Leu Lys Asn Leu Ser Ser His Trp Leu Met Lys
50 55 60
Ser Glu Pro Glu Ser Arg Leu Glu Lys Gly Val Asp Val Lys Phe
65 70 75
Ser Ile Glu Asp Leu Lys Ala Gln Pro Lys Gln Thr Thr Cys Trp
80 85 90
Asp Gly Val Arg Asn Tyr Gln Ala Arg Asn Phe Leu Arg Ala Met
95 100 105
Lys Leu Gly Glu Glu Ala Phe Phe Tyr His Ser Asn Cys Lys Glu
110 115 120
Pro Gly Ile Ala Gly Leu Met Lys Ile Val Lys Glu Ala Tyr Pro
125 130 135
Asp His Thr Gln Phe Glu Lys Asn Asn Pro His Tyr Asp Pro Ser
140 145 150
Ser Lys Glu Asp Asn Pro Lys Trp Ser Met Val Asp Val Gln Phe
155 160 165
Val Arg Met Met Lys Arg Phe Ile Pro Leu Ala Glu Leu Lys Ser
170 175 180
Tyr His Gln Ala His Lys Ala Thr Gly Gly Pro Leu Lys Asn Met
185 190 195
Val Leu Phe Thr Arg Gln Arg Leu Ser Ile Gln Pro Leu Thr Gln
200 205 210
Glu Glu Phe Asp Phe Val Leu Ser Leu Glu Glu Lys Glu Pro Ser
215 220 225

<210> 10
<211> 329
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1352789CD1

<400> 10
Met Ala Ala Gln Tyr Gly Ser Met Ser Phe Asn Pro Ser Thr Pro
1 5 10 15
Gly Ala Ser Tyr Gly Pro Gly Arg Gln Glu Pro Arg Asn Ser Gln
20 25 30
Leu Arg Ile Val Leu Val Gly Lys Thr Gly Ala Gly Lys Ser Ala
35 40 45
Thr Gly Asn Ser Ile Leu Gly Arg Lys Val Phe His Ser Gly Thr
50 55 60
Ala Ala Lys Ser Ile Thr Lys Lys Cys Glu Lys Arg Ser Ser Ser
65 70 75
Trp Lys Glu Thr Glu Leu Val Val Val Asp Thr Pro Gly Ile Phe
80 85 90

Asp Thr Glu Val Pro Asn Ala Glu Thr Ser Lys Glu Ile Ile Arg

```
                95                  100                 105
Cys Ile Leu Leu Thr Ser Pro Gly Pro His Ala Leu Leu Leu Val
                110                 115                 120
Val Pro Leu Gly Arg Tyr Thr Glu Glu Glu His Lys Ala Thr Glu
                125                 130                 135
Lys Ile Leu Lys Met Phe Gly Glu Arg Ala Arg Ser Phe Met Ile
                140                 145                 150
Leu Ile Phe Thr Arg Lys Asp Asp Leu Gly Asp Thr Asn Leu His
                155                 160                 165
Asp Tyr Leu Arg Glu Ala Pro Glu Asp Ile Gln Asp Leu Met Asp
                170                 175                 180
Ile Phe Gly Asp Arg Tyr Cys Ala Leu Asn Asn Lys Ala Thr Gly
                185                 190                 195
Ala Glu Gln Glu Ala Gln Arg Ala Gln Leu Leu Gly Leu Ile Gln
                200                 205                 210
Arg Val Val Arg Glu Asn Lys Glu Gly Cys Tyr Thr Asn Arg Met
                215                 220                 225
Tyr Gln Arg Ala Glu Glu Glu Ile Gln Lys Gln Thr Gln Ala Met
                230                 235                 240
Gln Glu Leu His Arg Val Glu Leu Glu Arg Glu Lys Ala Arg Ile
                245                 250                 255
Arg Glu Glu Tyr Glu Glu Lys Ile Arg Lys Leu Glu Asp Lys Val
                260                 265                 270
Glu Gln Glu Lys Arg Lys Lys Gln Met Glu Lys Lys Leu Ala Glu
                275                 280                 285
Gln Glu Ala His Tyr Ala Val Arg Gln Gln Arg Ala Arg Thr Glu
                290                 295                 300
Val Glu Ser Lys Asp Gly Ile Leu Glu Leu Ile Met Thr Ala Leu
                305                 310                 315
Gln Ile Ala Ser Phe Ile Leu Leu Arg Leu Phe Ala Glu Asp
                320                 325
```

<210> 11
<211> 237
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1666486CD1

<400> 11

```
Met Asp Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu
 1               5                  10                  15
Trp Leu Pro Gly Ala Lys Cys Asp Ile Leu Leu Thr Gln Ser Pro
                20                  25                  30
Ser Thr Val Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
                35                  40                  45
Arg Ala Thr Gln Ser Ile Gly Ser Trp Val Ala Trp Tyr Gln Gln
                50                  55                  60
Lys Pro Gly Lys Ala Pro Gln Leu Leu Ile Tyr Lys Ala Ser Ser
                65                  70                  75
Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
                80                  85                  90
Thr Glu Phe Thr Leu Ser Ile Asn Ser Leu Gln Pro Asp Asp Phe
                95                  100                 105
Ala Thr Tyr Phe Cys Gln Gln Tyr Asp Thr Tyr Pro Thr Trp Ser
                110                 115                 120
Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala
                125                 130                 135
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser
```

```
                    140                      145                      150
Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg
                    155                      160                      165
Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
                    170                      175                      180
Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
                    185                      190                      195
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
                    200                      205                      210
Lys His Lys Leu Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
                    215                      220                      225
Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
                    230                      235
```

```
<210> 12
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1706596CD1

<400> 12
Met Pro Trp Ala Leu Leu Leu Leu Thr Leu Leu Thr His Ser Ala
  1               5                  10                  15
Val Ser Val Val Gln Ala Gly Leu Thr Gln Pro Pro Ser Val Ser
                 20                  25                  30
Arg Ala Leu Arg Gln Thr Ala Thr Leu Thr Cys Thr Gly Asn Asn
                 35                  40                  45
Asn Asn Val Gly Asn Gln Gly Ala Ala Trp Leu Gln Gln His Gln
                 50                  55                  60
Gly His Pro Pro Lys Leu Leu Ser Tyr Arg Asn Asn Asn Arg Pro
                 65                  70                  75
Ser Gly Ile Ser Glu Arg Phe Ser Ala Ser Arg Ser Arg Asn Thr
                 80                  85                  90
Ala Ser Leu Thr Ile Thr Gly Leu Gln Pro Glu Asp Glu Ala Asp
                 95                 100                 105
Tyr Tyr Cys Ser Val Trp Asp Ser Ser Leu Ser Ala Trp Val Phe
                110                 115                 120
Gly Gly Gly Thr Lys Leu Thr Val Leu Ser Gln Pro Lys Ala Ala
                125                 130                 135
Pro Ser Val Thr Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala
                140                 145                 150
Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly
                155                 160                 165
Ala Val Thr Val Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala
                170                 175                 180
Gly Val Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys Tyr
                185                 190                 195
Ala Ala Ser Ser Tyr Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser
                200                 205                 210
His Arg Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr Val
                215                 220                 225
Glu Lys Thr Val Ala Pro Thr Glu Cys Ser
                230                 235

<210> 13
<211> 246
<212> PRT
```

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1890540CD1

<400> 13
Met Leu Trp Arg Gln Leu Ile Tyr Trp Gln Leu Leu Ala Leu Phe
1 5 10 15
Phe Leu Pro Phe Cys Leu Cys Gln Asp Glu Tyr Met Glu Ser Pro
20 25 30
Gln Thr Gly Gly Leu Pro Pro Asp Cys Ser Lys Cys Cys His Gly
35 40 45
Asp Tyr Ser Phe Arg Gly Tyr Gln Gly Pro Pro Gly Pro Pro Gly
50 55 60
Pro Pro Gly Ile Pro Gly Asn His Gly Asn Asn Gly Asn Asn Gly
65 70 75
Ala Thr Gly His Glu Gly Ala Lys Gly Glu Lys Gly Asp Lys Gly
80 85 90
Asp Leu Gly Pro Arg Gly Glu Arg Gly Gln His Gly Pro Lys Gly
95 100 105
Glu Lys Gly Tyr Pro Gly Ile Pro Pro Glu Leu Gln Ile Ala Phe
110 115 120
Met Ala Ser Leu Ala Thr His Phe Ser Asn Gln Asn Ser Gly Ile
125 130 135
Ile Phe Ser Ser Val Glu Thr Asn Ile Gly Asn Phe Phe Asp Val
140 145 150
Met Thr Gly Arg Phe Gly Ala Pro Val Ser Gly Val Tyr Phe Phe
155 160 165
Thr Phe Ser Met Met Lys His Glu Asp Val Glu Glu Val Tyr Val
170 175 180
Tyr Leu Met His Asn Gly Asn Thr Val Phe Ser Met Tyr Ser Tyr
185 190 195
Glu Met Lys Gly Lys Ser Asp Thr Ser Ser Asn His Ala Val Leu
200 205 210
Lys Leu Ala Lys Gly Asp Glu Val Trp Leu Arg Met Gly Asn Gly
215 220 225
Ala Leu His Gly Asp His Gln Arg Phe Ser Thr Phe Ala Gly Phe
230 235 240
Leu Leu Phe Glu Thr Lys
245

<210> 14
<211> 322
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2774913CD1

<400> 14
Met Pro Val Thr Val Thr Arg Thr Thr Ile Thr Thr Thr Thr Thr
1 5 10 15
Ser Ser Ser Gly Leu Gly Ser Pro Met Ile Val Gly Ser Pro Arg
20 25 30
Ala Leu Thr Gln Pro Leu Gly Leu Leu Arg Leu Leu Gln Leu Val
35 40 45
Ser Thr Cys Val Ala Phe Ser Leu Val Ala Ser Val Gly Ala Trp
50 55 60
Thr Gly Ser Met Gly Asn Trp Ser Met Phe Thr Trp Cys Phe Cys

```
                 65                  70                  75
Phe Ser Val Thr Leu Ile Ile Leu Ile Val Glu Leu Cys Gly Leu
                 80                  85                  90
Gln Ala Arg Phe Pro Leu Ser Trp Arg Asn Phe Pro Ile Thr Phe
                 95                 100                 105
Ala Cys Tyr Ala Ala Leu Phe Cys Leu Ser Ala Ser Ile Ile Tyr
                110                 115                 120
Pro Thr Thr Tyr Val Gln Phe Leu Ser His Gly Arg Ser Arg Asp
                125                 130                 135
His Ala Ile Ala Ala Thr Phe Phe Ser Cys Ile Ala Cys Val Ala
                140                 145                 150
Tyr Ala Thr Glu Val Ala Trp Thr Arg Ala Arg Pro Gly Glu Ile
                155                 160                 165
Thr Gly Tyr Met Ala Thr Val Pro Gly Leu Leu Lys Val Leu Glu
                170                 175                 180
Thr Phe Val Ala Cys Ile Ile Phe Ala Phe Ile Ser Asp Pro Asn
                185                 190                 195
Leu Tyr Gln His Gln Pro Ala Leu Glu Trp Cys Val Ala Val Tyr
                200                 205                 210
Ala Ile Cys Phe Ile Leu Ala Ala Ile Ala Ile Leu Leu Asn Leu
                215                 220                 225
Gly Glu Cys Thr Asn Val Leu Pro Ile Pro Phe Pro Ser Phe Leu
                230                 235                 240
Ser Gly Leu Ala Leu Leu Ser Val Leu Leu Tyr Ala Thr Ala Leu
                245                 250                 255
Val Leu Trp Pro Leu Tyr Gln Phe Asp Glu Lys Tyr Gly Gly Gln
                260                 265                 270
Pro Arg Arg Ser Arg Asp Val Ser Cys Ser Arg Ser His Ala Tyr
                275                 280                 285
Tyr Val Cys Ala Trp Asp Arg Arg Leu Ala Val Ala Ile Leu Thr
                290                 295                 300
Ala Ile Asn Leu Leu Ala Tyr Val Ala Asp Leu Val His Ser Ala
                305                 310                 315
His Leu Val Phe Val Lys Val
                320
```

```
<210> 15
<211> 191
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 5571291CD1

<400> 15
Met Trp Arg His Glu Arg Ile Lys Lys Thr Ser Phe Ser Thr Thr
  1               5                  10                  15
Thr Leu Leu Pro Pro Ile Lys Val Leu Val Val Tyr Pro Ser Glu
                 20                  25                  30
Ile Cys Phe His His Thr Ile Cys Tyr Phe Thr Glu Phe Leu Gln
                 35                  40                  45
Asn His Cys Arg Ser Glu Val Ile Leu Glu Lys Trp Gln Lys Lys
                 50                  55                  60
Lys Ile Ala Glu Met Gly Pro Val Gln Trp Leu Ala Thr Gln Lys
                 65                  70                  75
Lys Ala Ala Asp Lys Val Val Phe Leu Leu Ser Asn Asp Val Asn
                 80                  85                  90

Ser Val Cys Asp Gly Thr Cys Gly Lys Ser Glu Gly Ser Pro Ser
                 95                 100                 105
```

```
Glu Asn Ser Gln Asp Leu Phe Pro Leu Ala Phe Asn Leu Phe Cys
                110                 115                 120
Ser Asp Leu Arg Ser Gln Ile His Leu His Lys Tyr Val Val Val
                125                 130                 135
Tyr Phe Arg Glu Ile Asp Thr Lys Asp Asp Tyr Asn Ala Leu Ser
                140                 145                 150
Val Cys Pro Lys Tyr His Leu Met Lys Asp Ala Thr Ala Phe Cys
                155                 160                 165
Ala Glu Leu Leu His Val Lys Gln Gln Val Ser Ala Gly Lys Arg
                170                 175                 180
Ser Gln Ala Cys His Asp Gly Cys Cys Ser Leu
                185                 190
```

```
<210> 16
<211> 2265
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2705028CB1

<400> 16
cagttcagct cgctcggcgc acccacgcct cgctgccccg cttcctgccc tcaacctggg   60
catgctcccc ccacccttcc ggccccccag aacccgcgcc atcccccgga gcctccccag  120
agctggccgc gcaggatggg cgccctcagg cccacgctgc tgccgccttc gctgccgctg  180
ctgctgctgc taatgctagg aatgggatgc tgggcccggg aggtgctggt ccccgagggg  240
cccttgtacc gcgtggctgg cacagctgtc tccatctcct gcaatgtgac cggctatgag  300
ggccctgccc agcagaactt cgagtggttc ctgtataggc ccgaggcccc agatactgca  360
ctgggcattg tcagtaccaa ggatacccag ttctcctatg ctgtcttcaa gtcccgagtg  420
gtggcgggtg aggtgcaggt gcagcgccta caaggtgatg ccgtggtgct caagattgcc  480
cgcctgcagg cccaggatgc cggcatttat gagtgccaca cccccctccac tgatacccgc  540
tacctgggca gctacagcgg caaggtggag ctgagagttc ttccagatgt cctccaggtg  600
tctgctgccc ccccagggcc ccgaggccgc caggccccaa cctcaccccc acgcatgacg  660
gtgcatgagg ggcaggagct ggcactgggc tgcctggcga ggacaagcac acagaagcac  720
acacacctgg cagtgtcctt tgggcgatct gtgcccgagg caccagttgg gcggtcaact  780
ctgcaggaag tggtgggaat ccggtcagac ttggccgtgg aggctggagc tccctatgct  840
gagcgattgg ctgcagggga gcttcgtctg ggcaaggaag ggaccgatcg gtaccgcatg  900
gtagtagggg gtgcccaggc aggggacgca ggcacctacc actgcactgc cgctgagtgg  960
attcaggatc ctgatggcag ctgggcccag attgcagaga aaagggccgt cctggcccac 1020
gtggatgtgc agacgctgtc cagccagctg gcagtgacag tggggcctgg tgaacgtcgg 1080
atcggcccag gggagccctt ggaactgctg tgcaatgtgt caggggcact tcccccagca 1140
ggccgtcatg ctgcatactc tgtaggttgg gagatggcac ctgcgggggc acctgggccc 1200
ggccgcctgg tagcccagct ggacacagag ggtgtgggca gcctgggccc tggctatgag 1260
ggccgacaca ttgccatgga gaaggtggca tccagaacat accggctacg gctagaggct 1320
gccaggcctg gtgatgcggg cacctaccgc tgcctcgcca aagcctatgt tcgagggtct 1380
gggacccggc ttcgtgaagc agccagtgcc cgttcccggc ctctccctgt acacgtgcgg 1440
gaggaaggtg tggtgctgga ggctgtggca tggctagcag gaggcacagt gtaccgcggg 1500
gagactgcct ccctgctgtg caacatctct gtgcggggtg gcccccccagg actgcggctg 1560
gccgccagct ggtgggtgga gcgaccagag gacggagagc tcagctctgt ccctgcccag 1620
ctggtgggtg gcgtaggcca ggatggtgtg gcagagctgg gagtccggcc tggaggaggc 1680
cctgtcagcg tagagctggt ggggccccga agccatcggc tgagactaca cagcttgggg 1740
cccgaggatg aaggcgtgta ccactgtgcc cccagcgcct gggtgcagca tgccgactac 1800
agctggtacc aggcgggcag tgcccgctca gggcctgtta cagtctaccc ctacatgcat 1860
gccctggaca ccctatttgt gcctctgctg gtgggtacag ggtggccct agtcactggt 1920
gccactgtcc ttggtaccat cacttgctgc ttcatgaaga ggcttcgaaa acggtgatcc 1980
cttactcccc aggtcttgca ggtgtcgact gtcttccggc ccagctccaa gccctcctct 2040
ggttgcctgg acaccctctc cctctgtcca ctcttccttt aatttatttg acctcccact 2100
acccagaatg ggagacgtgc ctccccttcc ccactccttc cctcccaagc ccctccctct 2160
ggccttctgt tcttgatctc ttagggatcc tatagggagg ccatttcctg tcctggaatt 2220
agtttttcta aaatgtgaat aaacttgttt tataaaaaaa aaaaa                 2265
```

<210> 17
<211> 1124
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2751129CB1

<400> 17
ccgcccccac tgcccagagc cagagggatg gtggtagtca cggggcggga gccagacagc 60
cgtcgtcagg acggtgccat gtccagctct gacgccgaag acgactttct ggagccggcc 120
acgccgacgg ccacgcaggc ggggcacgcg ctgcccctgc tgccacagga gtttcctgag 180
gttgttcccc ttaacatcgg aggggctcac ttcactacac gcctgtccac actgcggtgc 240
tacgaagaca ccatgttggc agccatgttc agtgggcggc actacatccc cacggactcc 300
gagggccggt acttcatcga ccgagatggc acacactttg gagatgtgct gaatttcctg 360
cgctcagggg acctcccacc cagggagcgt gttcgagctg tgtacaaaga ggcccagtac 420
tatgccatcg ggcccctcct ggagcagctg gagaacatgc agccactgaa gggcgagaag 480
gtgcgccaag cgtttctggg actcatgccc tattacaaag accacttgga gcggattgtg 540
gagatcgccg gctgcgtgcg gtccagcgga aggcccgctt tgccaagctc aaggtctgtg 600
tcttcaagga ggagatgccc atcacccccct atgagtgtcc gctcctcaac tccctgcgat 660
ttgagcggag tgagagtgac gggcagcttt ttgagcacca ctgtgaagtg gatgtgtctt 720
ttgggccctg ggaggctgtg gctgatgttt atgacctgct gcactgcctg gtcacggacc 780
tctcggccca gggtctcacc gtggaccacc agtgcatcgg ggtgtgtgac aagcacctcg 840
tgaaccacta ctactgcaag ctgcccatct atgagttcaa gatcacatgg tggtgagtag 900
ctccggtagg cgagagtctc atcagggagg atgtccacct tgcttggtgg ctctgggagt 960
aagattcctg aaggggctgc tgactgccca gaatcctgcg aagtgagaac agcatcctga 1020
agcaaagctc ccagggacag aagtggtagt caatttcctg actgcactaa ggtttggctc 1080
aggtttcggc atgagantca ttcgngtaac tggctttctc agga 1124

<210> 18
<211> 1082
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2818076CB1

<400> 18
gcccgcgcag tccgcgcagc ctcatcgcaa ctgggcccgc gcgcaggcct tacataggaa 60
gtccttctaa agagctgcct gccagctgcc cttccccaga tccccgaatat cctcctggcc 120
aggtggagca gagaacagtt cctcagctgg tcatgctgag ctcataccct gatggctgct 180
ccatgaggtc aagactgggt ctcctccctc ctccccttc accaatgcct ggtctcacgg 240
ggctagtttt gacccccacg ctatggcatc atcgacctcc ctcccagctc ctggctctcg 300
gcctaagaag cctctaggca agatggctga ctggttcagg cagaccctgc tgaagaagcc 360
caagaagagg cccaactccc cagaaagcac ctccagcgat gcttcacagc ctacctcaca 420
ggacagccca ctacccccaa gcctcagctc agtcacgtct cccagcctgc caccacaca 480
tgcgagtgac agtggcagta gtcgctggag caaagactat gacgtctgcg tgtgccacag 540
tgaggaagac ctggtggccg cccaagacct ggtctcctac ttggaaggca gcactgccag 600
cctgcgctgc ttcctgcaac tccgggatgc aaccccaggc ggcgctatag tgtccgagct 660
gtgccaggca ctgagcagta gtcactgccg ggtgctgctc atcacgccgg gcttccttca 720
ggacccctgg tgcaagtacc agatgctgca ggccctgacc gaggctccag ggccgaggg 780
ctgcaccatc cccctgctgt tgggcctcag cagagctgcc tacccacctg agctccgatt 840
catgtactac gtcgatggca ggggccctga tggtggcttt cgtcaagtca aagaagctgt 900
catgcgttgt aagctactac aggagggaga aggggaacgg gattcagcta cagtatctga 960
tctactttga cttttaggag acagccctgt agcctagtag ttcaaagcgc agcttctgga 1020
aaaggctgtc ggggtttgta tcctggctcc tgcacttatt aacccataaa aagtaacttg 1080
tg 1082

<210> 19
<211> 1180

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2907049CB1

<400> 19
```
cagacagatg cttcattcct gtatggggtg gtattcctgc catgggtcct gggcttctcc   60
actggatggc cctttgtctc cttggaacag gtcatgggga tgccatggtc atccagaacc  120
caagatacca ggttacccag tttggaaagc cagtgaccct gagttgttct cagactttga  180
accataacgt catgtactgg taccagcaga agtcaagtca ggccccaaag ctgctgttcc  240
actactatga caaagatttt aacaatgaag cagacacccc tgataacttc caatccagga  300
ggccgaacac ttctttctgc tttcttgaca tccgctcacc aggcctgggg gacgcagcca  360
tgtacctgtg tgccaccagc aaatacaggg acggggagct gtttttgga gaaggctcta  420
ggctgaccgt actggaggac ctgaaaaacg tgttcccacc cgaggtcgct gtgtttgagc  480
catcagaagc agagatctcc cacacccaaa aggccacact ggtatgcctg gccacaggct  540
tctaccccga ccacgtggag ctgagctggt gggtgaatgg gaaggaggtg cacagtgggg  600
tcagcacaga cccgcagccc ctcaaggagc agcccgccct caatgactcc agatactgcc  660
tgagcagccg cctgagggtc tcggccacct tctggcagaa cccccgcaac cacttccgct  720
gtcaagtcca gttctacggg ctctcggaga atgacgagtg gacccaggat agggccaaac  780
ctgtcaccca gatcgtcagc gccgaggcct ggggtagagc agactgtggc ttcacctccg  840
agtcttacca gcaaggggtc ctgtctgcca ccatcctcta tgagatcttg ctagggaagg  900
ccaccttgta tgccgtgctg gtcagtgccc tcgtgctgat ggccatggtc aagagaaagg  960
attccagagg ctagctccaa aaccatccca ggtcattctt catcctcacc caggattctc 1020
ctgtacctgc tcccaatctg tgttcctaaa agtgattctc actctgcttc tcatctccta 1080
cttacatgaa tacttctctc tttttctgt ttccctgaag attgagctcc caacccccaa 1140
gtacgaaata ggctaaacca ataaaaaatt gtgtgttgga                      1180
```

<210> 20
<211> 1307
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3402252CB1

<400> 20
```
gatccacctt gattaaatct cttgtcccca gccctcctgg tccccaatgg cagcagccgc   60
catcccagct ttgcttctgt gcctcccgct tctgttcctg ctgttcggct ggtcccgggc  120
taggcgagac gaccctcact ctctttgcta tgacatcacc gtcatcccta agttcagacc  180
tggaccacgg tggtgtgcgg ttcaaggcca ggtggatgaa aagacttttc ttcactatga  240
ctgtggcaac aagacagtca cacctgtcag tcccctgggg aagaaactaa atgtcacaac  300
ggcctggaaa gcacagaacc cagtactgag agaggtggtg gacatactta cagagcaact  360
gcgtgacatt cagctggaga attacacacc caaggaaccc ctcaccctgc aggcaaggat  420
gtcttgtgag cagaaagctg aaggacacag cagtggatct tggcagttca gtttcgatgg  480
gcagatcttc ctcctctttg actcagagaa gagaatgtgg acaacggttc atcctggagc  540
cagaaagatg aaagaaaagt gggagaatga caaggttgtg gccatgtcct tccattactt  600
ctcaatggga gactgtatag gatggcttga ggacttcttg atgggcatgg acagcaccct  660
ggagccaagt gcaggagcac cactcgccat gtcctcaggc acaacccaac tcagggccac  720
agccaccacc ctcatccttt gctgcctcct catcatcctc ccctgcttca tcctccctgg  780
catctgagga gagtccttta gagtgacagg ttaaagctga taccaaaagg ctcctgtgag  840
cacggtcttg atcaaactcg cccttctgtc tggccagctg cccacgacct acggtgtatg  900
tccagtggcc tccagcagat catgatgaca tcatggaccc aatagctcat tcactgcctt  960
gattcctttt gccaacaatt ttaccagcag ttatacctaa catattatgc aattttctct 1020
tggtgctacc tgatggaatt cctgcactta aagttctggc tgactaaaca agatatatca 1080
ttttctttct tctcttttg tttggaaaat caagtacttc tttgaatgat gatctctttc 1140
ttgcaaatga tattgtcagt aaaataatca cgttagactt cagacctctg ggattctttt 1200
ccgtgtcctg aaagagaatt tttaaattat ttaataagaa aaaatttata ttaatgattg 1260
tttcctttag taatttattg ttctgtactg atatttaaat aacgcat               1307
```

<210> 21
<211> 689
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3577142CB1

<400> 21
cggcggcatc tggggttcca ggcgggcagc agctgcaggc tgaccttgca gcttggcgga 60
atggactggc ctcacaacct gctgtttctt cttaccattt ccatcttcct ggggctgggc 120
cagcccagga gccccaaaag caagaggaag ggcaagggc ggcctgggcc cctggtccct 180
ggcccctcacc aggtgccact ggacctggtg tcacggatga aaccgtatgc ccgcatggag 240
gagtatgaga ggaacatcga ggagatggtg gcccagctga ggaacagctc agagctggcc 300
cagagaaagt gtgaggtcaa cttgcagctg tggatgtcca acaagaggag cctgtctccc 360
tggggctaca gcatcaacca cgaccccagc cgtatccccg tggacctgcc ggaggcacgg 420
tgcctgtgtc tgggctgtgt gaacccttc accatgcagg aggaccgcag catggtgagc 480
gtgccggtgt tcagccaggt tcctgtgcgc cgccgcctct gcccgccacc gccccgcaca 540
gggccttgcc gccagcgcgc agtcatggag accatcgctg tgggctgcac ctgcatcttc 600
tgaatcacct ggcccagaag ccaggccagc agcccgagac catcctcctt gcacctttgt 660
gccaagaaag gcctatgaaa agtaaacac 689

<210> 22
<211> 818
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3725986CB1

<400> 22
aagggcaaga gccagctttc tctttccttg ctactttggc cagcagttcc agtgaagaga 60
tggagtctca gctgcaggaa cgtgtggagt cttcccgccg agccgtgtcc cagattgtga 120
ctgtttatga taaattgcaa gaaaaagtgg agctcttatc ccggaagcta aacagtggag 180
ataatctgat agtggaggaa gcagtgcagg agctgaactc tttcctcgca caggagaata 240
tgaggctaca ggaattgaca gatcttcttc aggaaaagca tcgcaccatg tctcaggagt 300
tctccaagtt gcagagtaaa gtggagacag ccgaatcacg agtgtctgtc ctggagtcca 360
tgattgatga cctgcagtgg gatattgaca aaattcgaaa gagggaacag cgactcaacc 420
gacacttagc agaagtccta gaacgggtga attccaaagg ttataaggtg tatggagcgg 480
ggagcagtct gtatggcggc acaatcacta tcaatgctcg gaagtttgag gaaatgaatg 540
cagagcttga ggagaacaaa gagttggctc agaaccgtct ctgtgagctg gagaaacttc 600
ggcaagactt tgaggaggtc actacacaaa atgaaaagct gaaggtacga acgcatccct 660
gaagggcagt aaaatcagac gttctgctga tcaactcacg tatatacata gttgtgaatc 720
tgcgtattca tgagggataa gaaaaatgta gacaaaatcc aacatccttt tatgataaaa 780
ctcttaacaa attaggtgta aagaagtgta cctcaaca 818

<210> 23
<211> 899
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3799011CB1

<400> 23
gccttggagg cgtgactttc ccctcgggtc caggtagggc ctggagctgc tgcaagtgcc 60
gcctgtgctg gggaagggac catgtggctg ccttgggctc tgttgcttct ctgggtccca 120
ggatgttttg ctctgagcaa atgcaggacc gtggcgggcc ccgtgggggg atccctgagt 180
gtgcagtgtc cctatgagaa ggaacacagg accctcaaca aatactggtg cagaccacca 240

cagattttcc tatgtgacaa gattgtggag accaaagggt cagcaggaaa aaggaacggc 300
cgagtgtcca tcagggacag tcctgcaaac ctcagcttca cagtgaccct ggagaatctc 360
acagaggagg atgcaggcac ctactggtgt ggggtggata caccatggct ccgagacttt 420
catgatcccg ttgtcgaggt tgaggtgtcc gtgttcccgg catcaacgtc aatgacacct 480
gcaagtatca ctgcggccaa gacctcaaca atcacaactg catttccacc tgtatcatcc 540
actaccctgt ttgcagtggg tgccacccac agtgccagca tccaggagga aactgaggag 600
gtggtgaact cacagctccc gctgactcct ctccctgctg gcattgttgc tgcttctgtt 660
ggtgggggcc tccctgctag cctggaggat gtttcagaaa tggatcaaag ctggtgacca 720
ttcagagctg tcccagaacc ccaagcaggc tgccacgcag agtgagctgc actacgcaaa 780
tctggagctg ctgatgtggc ctctgcagga aaagccagca ccaccaaggg aggtggaggt 840
ggaatacagc actgtggcct cccccaggga agaacttcac tatgcctcgg tggtgtttg 899

<210> 24
<211> 953
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3887384CB1

<400> 24
cgcgggggtc gcgctgcaca gcctgcggcg cagcggaggc ggaccgcagt cgagtctgca 60
gagtgttgga atcgtccgca ctaaagtccc ctgcagcgtg accatgtcga gaccccggaa 120
gaggctggct gggacttctg gttcagacaa gggactatca ggaaaacgca ccaaaactga 180
gaactcaggt gaggcattag ctaaagtgga ggactccaac cctcagaaga cttcagccac 240
taaaaactgt ttgaagaatc taagcagcca ctggctgatg aagtcagagc cagagagccg 300
cctagagaaa ggtgtagatg tgaagttcag cattgaggat ctcaaagcac agcccaaaca 360
gacaacatgc tgggatggtg ttcgtaacta ccaggctcgg aacttcctta gagccatgaa 420
gctgggagaa gaagccttct tctaccatag caactgcaaa gagccaggca tcgcaggact 480
catgaagatc gtgaaagagg cttacccaga ccacacacag tttgagaaaa acaatcccca 540
ttatgaccca tctagcaaag aggacaaccc taagtggtcc atggtggatg tacagtttgt 600
tcggatgatg aaacgtttca ttcccctggc tgagctcaaa tcctatcatc aagctcacaa 660
agctactggt ggccccttaa aaaatatggt tctcttcact cgccagagat tatcaatcca 720
gcccctgacc caggaagagt ttgattttgt tttgagcctg gaggaaaagg aaccaagtta 780
actgagatac tgctgctgga atgggcgaga cattgctgca aagaagtcaa gctttttca 840
gacaaaaggt gtgaggggc ttgcttggta tgcttacctg ggcttgtgta cctcagtggt 900
ttttgtgtac ttttttcaat aaaatatcaa agttgaagaa aaaaaaaaa aaa 953

<210> 25
<211> 1979
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1352789CB1

<400> 25
ttctcaacat cctggcttag tattgtgtgc aaaatcagag aggggtgcaa gatcctgatt 60
tttcaggagt tcaagcgaca atggcagccc aatacggcag tatgagcttc aaccccagca 120
caccaggggc cagttatggg cctggaaggc aagagcccag aaattcccaa ttgagaattg 180
tgttagtggg taaaaccgga gcaggaaaaa gtgcaacagg aaacagcatc cttggccgga 240
aagtgtttca ttctggcact gcagcaaaat ccattaccaa gaagtgtgag aaacgcagca 300
gctcatggaa ggaaacagaa cttgtcgtag ttgacacacc aggcattttc gacacagagg 360
tgcccaatgc tgaaacgtcc aaggagatta ttcgctgcat tcttctgacc tccccagggc 420
ctcatgctct gcttctggtg gttccactgg gccgttacac tgaggaagag cacaaagcca 480
cagagaagat cctgaaaatg tttggagaga gggctagaag tttcatgatt ctcatattca 540
cccggaaaga tgacttaggt gacaccaatt tgcatgacta cttaagggaa gctccagaag 600
acattcaaga cttgatggac attttcggtg accgctactg tgcgttaaac aacaaggcaa 660

caggcgctga gcaggaggcc cagagggcac agttgctggg cctgatccag cgcgtggtga 720
gggagaacaa ggaaggctgc tacactaata ggatgtacca aagggcggag gaggagatcc 780
agaagcaaac acaagcaatg caagaactcc acagagtgga gctggagaga gagaaagcgc 840
ggataagaga ggagtatgaa gagaaaatca gaaagctgga agataaagtg gagcaggaaa 900
agagaaagaa gcaaatggag aagaaactag cagaacagga ggctcactat gctgtaaggc 960
agcaaagggc aagaacggaa gtggagagta aggatgggat acttgaatta atcatgacag 1020
cgttacagat tgcttccttt attttgttac gtctgttcgc ggaagattaa acttaatgaa 1080
aatctgtttg tattttctgc atattctctg gcaaccttgc cccatactta cttatttagc 1140
atagtcgagt gctctagttt ctgtctctca ggcactcgta actaaggacc accattggcc 1200
attggtagat gtttgattga cttaacaaga gagggacaaa ttttcaattt gtgaaactcc 1260
aaagcagaaa gtattggtgc ttgctacctt gtgaattctt ccttagacat gcagagaaaa 1320
tgtatgcaag agaccaaaaa gatggctcca agctatgtca tgttacctgt aataaaatct 1380
tttcttctag attctttcta tgttggcaga taatctcccc ttgtagcttc cactcactta 1440
ttcttgcatt cagagtcaca atgatcatct tacccatgtg gtttttgaga aagaaagatc 1500
aattctttgt ttgcagtagg taatcttaga gatggagatg attgtagaat tattcctaga 1560
tgagtgtcaa tttatttaat tccattgtca tataaggagt caaattgttt cttatcattt 1620
gttcattgaa gaacagagac ctgtctggaa aatcgatctc tacaaattca attaaataat 1680
gatccccaaa tgctgaaaaa gtgaaataca gcaattcaac agataataga gcaatgttta 1740
gtatattcag ctgtatctgt agaaactctt tgacgaacct caatttaacc aatttgatga 1800
atacccagtt ctcttctttt ctagagaaag atagttgcaa cctcacctcc ctcactcaac 1860
actttgaata cttattgttt ggcaggtcat ccacacactt ctgcccccac tgcattgaat 1920
tttttgctta tgttgtttat aataaaactt ttcaattatc tcataaaaaa aaaaaaaaa 1979

<210> 26
<211> 923
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1666486CT1

<400> 26
cccagtcagg acacagcatg gacatgaggg tccccgctca gctcctgggg ctcctgctgc 60
tctggctccc aggtgccaaa tgtgacatcc tgctgaccca gtctccttcc accgtctctg 120
catctgtggg cgacagagtc accattactt gccgggccac tcagagtatt ggtagctggg 180
tggcctggta tcagcaaaaa ccagggaaag cccctcagct cctgatctat aaggcgtcca 240
gtttagaaag tggagtccca tcaaggttta gtggcagtgg atctgggaca gaattcactc 300
tcagcatcaa cagcctgcag cctgatgatt ttgcaactta tttctgccag cagtatgaca 360
cttaccccac gtggtcgttc ggccagggga ccaagctgga gatcaaacga actgtggctg 420
caccatctgt cttcatcttc ccgccatctg atgagcagtt gaaatctgga actgcctctg 480
ttgtgtgcct gctgaataac ttctatccca gagaggccaa agtacagtgg aaggtggata 540
acgccctcca atcgggtaac tcccaggaga gtgtcacaga gcaggacagc aaggacagca 600
cctacagcct cagcagcacc ctgacgctga gcaaagcaga ctacgagaaa cacaaactct 660
acgcctgcga agtcacccat cagggcctga gctcgcccgt cacaaagagc ttcaacaggg 720
gagagtgtta gagggagaag tgcccccacc tgctcctcag ttccagcctg accccctccc 780
atcctttggc ctctgaccct tttccacag gggacctacc cctattgcgg tcctccagct 840
catctttcac ctcacccccc tcctcctcct tggctttaat tatgctaatg ttggaggaga 900
atgaataaat aaagtgaatc ttc 923

<210> 27
<211> 888
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1706596CB1

<400> 27

cagcaagcat agtgggaatc tgcaccatgc cctgggctct gctcctcctg accctcctca 60
ctcactctgc agtgtcagtg gtccaggcag ggctgactca gccaccctcg gtgtccaggg 120
ccttgagaca gaccgccaca ctcacctgca ctgggaacaa caacaatgtt ggcaaccaag 180
gagcagcttg gctgcagcag caccagggcc accctcccaa actcctgtcc tacaggaata 240
acaaccggcc ctcagggatc tcagagagat tctctgcatc caggtcgaga aacacagcct 300
ccctgaccat tactggactc cagcctgagg acgaggctga ctattactgc tcagtatggg 360
acagcagcct cagtgcttgg gtgttcggcg gagggaccaa gctgaccgtc ctaagtcagc 420
ccaaggctgc cccctcggtc actctgttcc caccctcctc tgaggagctt caagccaaca 480
aggccacact ggtgtgtctc ataagtgact tctacccggg agccgtgaca gtggcctgga 540
aggcagatag cagccccgtc aaggcgggag tggagaccac cacaccctcc aaacaaagca 600
acaacaagta cgcggccagc agctacctga gcctgacgcc tgagcagtgg aagtcccaca 660
gaagctacag ctgccaggtc acgcatgaag ggagcaccgt ggagaagaca gtggccccta 720
cagaatgttc ataggttctc aaccctcacc ccccaccacg ggagactaga gctgcaggat 780
cccaggggag gggtctctcc tcccaccccа aggcatcaag cccttctccc tgcactcaat 840
aaaccctcaa taaatattct cattgtcaat cagaaaaaaa aaaaaaa 888

<210> 28
<211> 1760
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1890540CB1

<400> 28
ccaaagtata aaaccactcc gccgctgcag ctctcagctc cagtcctggc atctgcccga 60
ggagaccacg ctcctggagc tctgctgtct tctcagggag actctgaggc tctgttgaga 120
atcatgcttt ggaggcagct catctattgg caactgctgg ctttgttttt cctccctttt 180
tgcctgtgtc aagatgaata catggagtct ccacaaaccg gaggactacc cccagactgc 240
agtaagtgtt gtcatggaga ctacagcttt cgaggctacc aaggcccccc tgggccaccg 300
ggccctcctg gcattccagg aaaccatgga aacaatggca acaatggagc cactggtcat 360
gaaggagcca aaggtgagaa gggcgacaaa ggtgacctgg ggcctcgagg ggagcggggg 420
cagcatggcc ccaaaggaga gaagggctac ccggggattc caccagaact tcagattgca 480
ttcatggctt ctctggcaac ccacttcagc aatcagaaca gtgggattat cttcagcagt 540
gttgagacca acattggaaa cttctttgat gtcatgactg gtagatttgg ggccccagta 600
tcaggtgtgt atttcttcac cttcagcatg atgaagcatg aggatgttga ggaagtgtat 660
gtgtacctta tgcacaatgg caacacagtc ttcagcatgt acagctatga aatgaagggc 720
aaatcagata catccagcaa tcatgctgtg ctgaagctag ccaaagggga tgaggtttgg 780
ctgcgaatgg gcaatggcgc tctccatggg gaccaccaac gcttctccac ctttgcagga 840
ttcctgctct ttgaaactaa gtaaatatat gactagaata gctccacttt ggggaagact 900
tgtagctgag ctgatttgtt acgatctgag gaacattaaa gttgagggtt ttacattgct 960
gtattcaaaa aattattggt tgcaatgttg ttcacgctac aggtacacca ataatgttgg 1020
acaattcagg ggctcagaag aatcaaccac aaaatagtct tctcagatga ccttgactaa 1080
tatactcagc atctttatca ctctttcctt ggcacctaaa agataattct cctctgacgc 1140
aggttggaaa tatttttttc tatcacagaa gtcatttgca aagaattttg actactctgc 1200
ttttaattta ataccagttt tcaggaaccc ctgaagtttt aagttcatta ttctttataa 1260
catttgagag aatcagatgt agtgatatga cagggctggg gcaagaacag gggcactagc 1320
tgccttatta gctaatttag tgccctccgt gttcagctta gcctttgacc ctttcctttt 1380
gatccacaaa atacattaaa actctgaatt cacatacaat gctattttaa agtcaataga 1440
ttttagctat aaagtgcttg accagtaatg tggttgtaat tttgtgtatg ttcccccaca 1500
tcgcccccaa cttcggatgt ggggtcagga ggttgaggtt cactattaac aaatgtcata 1560
aatatctcat agaggtacag tgccaataga tattcaaatg ttgcatgttg accagaggga 1620
ttttatatct gaagaacata cactattaat aaatacctta gagaaagatt ttgacctggc 1680
tttagataaa actgtggcaa gaaaaatgta atgagcaata tatggaaata aacacacctt 1740
tgttaaagat aaaaaaaaaa 1760

<210> 29
<211> 2015
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2774913CB1

<400> 29
gcttcagccg cagtcgccac tggctgcctg aggtgctctt acagcctgtt ccaagtgtgg 60
cttaatccgt ctccaccacc agatctttct ccgtggattc ctctgctaag accgctgcca 120
tgccagtgac ggtaacccgc accaccatca caaccaccac gacgtcatct tcgggcctgg 180
ggtcccccat gatcgtgggg tcccctcggg ccctgacaca gcccctgggt ctccttcgcc 240
tgctgcagct ggtgtctacc tgcgtggcct tctcgctggt ggctagcgtg ggcgcctgga 300
cggggtccat gggcaactgg tccatgttca cctggtgctt ctgcttctcc gtgaccctga 360
tcatcctcat cgtggagctg tgcgggctcc aggcccgctt ccccctgtct tggcgcaact 420
tccccatcac cttcgcctgc tatgcggccc tcttctgcct ctcggcctcc atcatctacc 480
ccaccaccta tgtccagttc ctgtcccacg gccgttcgcg ggaccacgcc atcgccgcca 540
ccttcttctc ctgcatcgcg tgtgtggctt acgccaccga agtggcctgg acccgggccc 600
ggcccggcga gatcactggc tatatggcca ccgtacccgg gctgctgaag gtgctggaga 660
ccttcgttgc ctgcatcatc ttcgcgttca tcagcgaccc caacctgtac cagcaccagc 720
cggccctgga gtggtgcgtg gcggtgtacg ccatctgctt catcctagcg gccatcgcca 780
tcctgctgaa cctgggggag tgcaccaacg tgctacccat cccttcccc agcttcctgt 840
cggggctggc cttgctgtct gtcctcctct atgccaccgc ccttgttctc tggcccctct 900
accagttcga tgagaagtat ggcggccagc ctcggcgctc gagagatgta agctgcagcc 960
gcagccatgc ctactacgtg tgtgcctggg accgccgact ggctgtggcc atcctgacgg 1020
ccatcaacct actggcgtat gtggctgacc tggtgcactc tgcccacctg gtttttgtca 1080
aggtctaaga ctctcccaag aggctcccgt tccctctcca acctctttgt tcttcttgcc 1140
cgagttttct ttatggagta cttctttcct ccgcctttcc tctgttttcc tcttcctgtc 1200
tcccctccct cccacctttt tctttccttc ccaattcctt gcactctaac cagttcttgg 1260
atgcatcttc ttccttccct ttcctcttgc tgtttccttc ctgtgttgtt ttgttgccca 1320
catcctgttt tcaccccctga gctgtttctc tttttctttt ctttcttttt ttttttttt 1380
tttaagacgg attctcacca ctgtgctcca gcctgggggga cagagcgaga ctccatctca 1440
aaaaaaaaaa ggaatcggac gaagaaccac aggatgttga agacaactgt ctgaagtatt 1500
tgtgagggac agcgatgtgg ccctctgtgt taagaataac gtgtcctgct ttggcagaga 1560
gaagaaaata gccactgccc gctttcaagg caagatcgac cttttctgtt ttgttttgtt 1620
tttctttctt tttcctggcc atgaggacaa aaattactga gtggccctta aagagggaag 1680
tttgttttca gctgttctct tttgcccgta ggtgggaggg tggggattgc tgcgtcctag 1740
ctagaggaat ggctttgctt gaatgtgtag tgcacacgca cgggtgtttc tgtgtgctag 1800
ttgcttcttg ctgctgcttc ctgcttgtct gggactcaca tacataacgt gatatatata 1860
tatatatata aatgtataaa tatatatttt atttttttt aaatccttgg agcttctggt 1920
tcctatcagt tcctgttgtt aatcgtagaa ccgttgtccc ttcccccatt cccgtatcca 1980
tcatgttctt tttcttttaa atatcaatat aaaaa 2015

<210> 30
<211> 2080
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 5571291CB1

<400> 30
gctgcagtga accgagattg cgccactgca ctccagccta ggcaacaaag cgagactctg 60
tctcaaaact aaaaaataat aaaaaataaa taaacctcta taaagtatac caagtcttag 120
tttttaaatt aagagataag tgtggatttg ttttccaaag gtgaataagc tttgtttttt 180
ccagacaaaa gcaagccggg aggctggctg cctctcctcc tgctgtctct gctggtggcc 240
acatgggtgc tggtggcagg gatctatcta atgtggaggc acgaaaggat caagaagact 300
tcctttttcta ccaccacact actgcccccc attaaggttc ttgtggttta cccatctgaa 360
atatgtttcc atcacacaat ttgttacttc actgaatttc ttcaaaacca ttgcagaagt 420
gaggtcatcc ttgaaaagtg gcagaaaaag aaaatagcag agatgggtcc agtgcagtgg 480
cttgccactc aaaagaaggc agcagacaaa gtcgtcttcc ttctttccaa tgacgtcaac 540
agtgtgtgcg atggtacctg tggcaagagc gagggcagtc ccagtgagaa ctctcaagac 600
ctcttccccc ttgcctttaa ccttttctgc agtgatctaa gaagccagat tcatctgcac 660
aaatacgtgg tggtctactt tagagagatt gatacaaaag acgattacaa tgctctcagt 720
gtctgcccca agtaccacct catgaaggat gccactgctt tctgtgcaga acttctccat 780

```
gtcaagcagc aggtgtcagc aggaaaaaga tcacaagcct gccacgatgg ctgctgctcc 840
ttgtagccca cccatgagaa gcaagagacc ttaaaggctt cctatcccac caattacagg 900
gaaaaaacgt gtgatgatcc tgaagcttac tatgcagcct acaaacagcc ttagtaatta 960
aaacatttta taccaataaa attttcaaat attgctaact aatgtagcat taactaacga 1020
ttggaaacta catttacaac ttcaaagctg ttttatacat agaaatcaat tacagtttta 1080
attgaaaact ataaccattt tgataatgca acaataaagc atcttcagcc aaacatctag 1140
tcttccatag accatgcatt gcagtgtacc cagaactgtt tagctaatat tctatgttta 1200
attaatgaat actaactcta agaacccctc actgattcac tcaatagcat cttaagtgaa 1260
aaaccttcta ttacatgcaa aaaatcattg tttttaagat aacaaaagta gggaataaac 1320
aagctgaacc cacttttact ggaccaaatg atctattata tgtgtaacca cttgtatgat 1380
ttggtatttg cataagacct tccctctaca aactagattc atatcttgat tcttgtacag 1440
gtgcctttta acatgaacaa caaaataccc acaaacttgt ctacttttgc ctaaagttac 1500
ctattagagg tcactgtcag agttctcagt ttcttagtta ctatttaact tttcatgttc 1560
aaaatgaaaa taattcttaa gttgaaagcc ctcttgaagt aacctttta taaatgagtt 1620
attataatgg tttacttaaa taaaaaacag gggtgggtgc agtggctcat gcctccaatc 1680
ccagcacttt ggcaaggcca aggcaaaagg atcgctcaag accaggctac gtcacaaagc 1740
gagacctcca tctctacaaa agatttaaaa aattagctga gtgtgatggt gtaagcctgt 1800
ggtcccagct actagggagg ctgagatggg aggatcactt gagccctgga ggtcaagggt 1860
gcagtaaacg gtgattgtgc cactgcactc catcctgggt gagagcagac cctgtctaaa 1920
acaaacaaac gaaaaaaccc ccacagaatg acagaacata aaagatgcac attttgtctt 1980
cccacttttt tactcttcta aaggcatctt tttttaaatt ttcttaaatt tttttttttt 2040
tgagacagag gttcacccgg tcacccaggc tggaggtgcg                       2080
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 October 2000 (12.10.2000)

PCT

(10) International Publication Number
WO 00/60080 A3

(51) International Patent Classification[7]: C12N 15/12, 5/10, C07K 14/47, 14/705, 16/18, 16/28, A61K 38/17, C12Q 1/68, G01N 33/50

(21) International Application Number: PCT/US00/09072

(22) International Filing Date: 4 April 2000 (04.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/127,852 5 April 1999 (05.04.1999) US
60/132,647 5 May 1999 (05.05.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:
US 60/127,852 (CIP)
Filed on 5 April 1999 (05.04.1999)
US 60/132,647 (CIP)
Filed on 5 May 1999 (05.05.1999)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

(88) Date of publication of the international search report: 25 January 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/60080 A3

(54) Title: MOLECULES OF THE IMMUNE SYSTEM

(57) Abstract: The invention provides human immune system molecules (IMOL) and polynucleotides which identify and encode IMOL. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of IMOL.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/09072

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N5/10 C07K14/47 C07K14/705 C07K16/18
C07K16/28 A61K38/17 C12Q1/68 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K A61K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, BIOSIS, STRAND, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] Accession No. AI336860, 31 December 1998 (1998-12-31) STRAUSBERG R.: "EST; H. sapiens cDNA clone IMAGE:2009429 similar to leukocyte surface protein" XP002143583 nt 119-307 encode aa 148-289 of seq. ID 1 ---	1-17,20, 23
A	RUEGG C.L. ET AL.: "V7, a novel leukocyte surface protein that participates in T cell activation. II. Molecular cloning and characterization of the V7 gene." J. IMMUNOL., vol. 154, 1995, pages 4434-4443, XP002143613 the whole document ---	1-17,20, 23

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 July 2000	08.11.00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Galli, I

2

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/09072

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 27114 A (SCHERING CORP) 25 June 1998 (1998-06-25) abstract claims 1-10	1-17,20, 23

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

International application No.
PCT/US 00/09072

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (C ntinuation f it m 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Claims 16,19,22 are directed to methods of treatment of the human/animal body. The search, if at all possible, has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 18,19,21,22
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-23, partially.

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 18,19,21,22

Claims 18,19,21,22 relate to agonists and antagonists of the human immune system molecule claimed, but do not offer a true technical characterization thereof. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and the subject-matter is not adequately disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims, the wording of which is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1-23) - partial

   An isolated polypeptide comprising the amino acid sequence of seq. ID 1.

   Corresponding nucleic acids (Seq. ID 16), vectors, recombinant host cells, transgenic organisms, antibodies, pharmaceutical compositions, detection methods, methods to evaluate agonists/antagonists/regulators, therapeutic applications.

2. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 2,17

3. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 3,18

4. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 4,19

5. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 5,20

6. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 6,21

7. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 7,22

8. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 8,23

9. Claims: (1-23) - partly

   Idem as subject-matter 1, but limited to seq. IDs 9,24

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 10,25

11. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 11,26

12. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 12,27

13. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 13,28

14. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 14,29

15. Claims: (1-23) - partly

    Idem as subject-matter 1, but limited to seq. IDs 15,30

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ational Application No
PCT/US 00/09072

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9827114 A	25-06-1998	AU 5705898 A	15-07-1998

Form PCT/ISA/210 (patent family annex) (July 1992)